Simple, efficient and robust techniques for automatic multi-objective function parameterisation: case studies of local and global optimisation using APSIM

Matthew Tom Harrison[A]*, Pier Paolo Roggero[B] and Laura Zavattaro[C]

[A] Tasmanian Institute of Agriculture, University of Tasmania, 16-20 Mooreville Rd, Burnie, Tasmania, 7320

[B] Desertification Research Centre and Department of Agricultural Sciences, University of Sassari, viale Italia 39, 07100 Sassari, Italy

[C]Dept. of Agricultural Forest and Food Sciences, University of Turin, largo P. Braccini, 2, 10095 Grugliasco, Italy

*Corresponding author: Matthew.Harrison@utas.edu.au

** Published in *Environmental Modelling & Software*, March 2019.

Highlights

- Automated parameterisation techniques are explored using the freeware PEST
- Optimisation was fastest when the Gauss-Levenberg-Marquardt algorithm was employed
- Tikhonov regularisation with SVD or LSQR significantly improved model calibration
- CMAES resulted in the best fit but required the longest optimisation times
- Log transformation of parameters generally improved calibration quality

Abstract

Several techniques for automatic parameterisation are explored using the software PEST. We parameterised the biophysical systems model APSIM with measurements from a maize cropping experiment with the objective of finding algorithms that resulted in the least distance between modelled and measured data (ϕ) in the shortest possible time. APSIM parameters were optimised using a weighted least-squares approach that minimised the value of ϕ. Optimisation techniques included the Gauss-Marquardt-Levenberg (GML) algorithm, singular value decomposition (SVD), least squares with QR decomposition (LSQR), Tikhonov regularisation, and covariance matrix adaptation-evolution strategy (CMAES).

In general, CMAES with log transformed APSIM parameters and larger population size resulted in the lowest ϕ, but this approach required significantly longer to converge compared with other optimisation algorithms. Regularisation treatments with log transformed parameters also resulted in low ϕ values when combined with SVD or LSQR; LSQR treatments with no regularisation tended to converge earliest.

In addition to an analysis of several PEST algorithms, this study provides a narrative on how methodologies presented here could be generalised and applied to other models.

Keywords

CPU time; genetic algorithm; inverse modelling; optimization; parameterization; regularization

Software availability

- APSIM version 7.8, programmed in C#.NET and VB.NET, freely available subject to user licencing at http://www.apsim.info/Products/Downloads.aspx
- PEST version 14.2, programmed in FORTRAN, freely available at http://www.pesthomepage.org/ , Contact address Watermark Numerical Computing, 336 Cliveden Avenue, Corinda 4075, Australia. Telephone 07 3379 1664; Email address johndoherty@ozemail.com.au

Introduction

Parameterisation (or parameterization, or calibration) is the process of adjusting the parameters of a mathematical model to improve the agreement or fit between model outputs and the observed or measured data (Wallach et al., 2014). In modelling specific agricultural scenarios, such as crop production in a specific location, models typically first require parameterisation, wherein their outputs conform with measured data. An extensive survey of methods used to parameterise crop models found that nearly half the 211 respondents used trial-and-error to search for the best-fit parameters (Seidel et al., 2018), such as that applied by Harrison et al. (2012). However, manual calibration techniques have the disadvantage of allowing only a small number of parameters to be calibrated, with a large amount of parameter combinations remaining uninvestigated. Indeed, often such approaches involve less than 10 parameters (Seidel et al., 2018). Recently Holzworth et al. (2015) emphasised the need for a more objective and reproducible calibration and validation methodology as a way forward for models in a growing agricultural domain, suggesting that the

availability of a reproducible calibration methodology helps simplify model calibration documentation in an industry where model documentation has been a long-standing issue (Holzworth et al., 2015; Sexton et al., 2016).

Automated approaches aim to optimise model parameters through programmed directives that assess sequential changes in the objective function (ϕ), often calculated as the cumulative squared difference between the observed data and corresponding modelled data. Auto-parameterisation methods typically begin with initial parameter sets based on expert knowledge (prior information) and continue to be iteratively upgraded and thus form new parameter vectors until specified termination criteria are reached. In many algorithms, such criteria are based on the rate of change in parameters and ϕ over consecutive iterations. Optimisation problems of this kind may involve several variables, such that ϕ is comprised by multiple components. Although automated approaches for model parameterisation have existed for some time (Lacroix et al., 2002; Samanta and Mackay, 2003; Sequeira et al., 1994), there are few studies that have examined whether auto-parameterisation can be used to calibrate dynamic programs such as APSIM (Keating et al., 2003) (but see notable exceptions by Akponikpè et al. (2010), Chen et al. (2016) and Sexton et al. (2016)). The limitation of past work performing optimisation of agricultural models may be because the objective functions formed by using such models often have discontinuities that make it difficult to use gradient-based minimisation methods (Buis et al., 2011). A common approach, such as that adopted in the crop model OptimiSTICS, is to use the Nelder–Mead simplex algorithm, which is adapted to non-smooth functions because the search of the optimum is not based on the computation of the function's gradient. As the Nelder–Mead simplex is a local optimisation method, OptimiSTICS automatically repeats the minimisation with several different starting parameter values to minimise the risk of converging to a local minimum. However, since this approach requires the user to specify the number of starting points, as well as the starting values in the options file (Buis et al., 2011), it does not guarantee convergence on the global minimum of the response surface. Other approaches, such as generalised likelihood uncertainty estimation (GLUE) (Chisanga et al., 2015; Sexton et al., 2016) and Bayesian parameter estimation (Wallach et al., 2012), yield parameter estimates that are strongly dependent on the choice of likelihood function and the method of combining likelihood values (Seidel et al., 2018). Nonetheless, work by Sexton et al. (2016) showed that both GLUE and Markov Chain Monte Carlo (MCMC) calibrations resulted in accurate simulations of biomass and yield in the crop model APSIM-Sugar.

Given the importance of phenology in developing new cultivars, many previous agricultural model optimisation studies have concentrated on crop phenological parameters. Indeed, programs such as GENCALC (Hunt et al., 1993), and more recently the GLUE optimisation program (Jones et al., 2011) have been included in the latest release of the DSSAT model (Jones et al., 2003). These tools allow DSSAT users to calibrate phenological parameters (and other crop parameters) of photoperiod-sensitive plants from few observations (e.g. Marin et al., 2011). The NAG (1983) is another dedicated optimiser that has been used to estimate phenological parameters for APSIM crops (Carberry et al., 2001; Farré et al., 2004; Turpin et al., 2003). However, restriction to only phenological parameters has meant that such studies have not simultaneously optimised other parameters, such as those related to soil water content or nitrous oxide emissions. Indeed, some studies have extracted model algorithms and programmed them in separate programs (e.g. Archontoulis et al., 2014 used R to program phenological equations in APSIM), which increases the risk that dynamic interactions between APSIM modules (such as SoilN, Plant and SurfaceOM) are not captured while parameter optimisation takes place. In this study, we aimed to conduct simultaneous optimisation of several model parameters, not just phenology.

The model-independent Parameter ESTimation software PEST (Doherty, 2016a) has been used successfully with studies of soil biogeochemical models (Necpálová et al., 2015), fractured porous media (Finsterle and Zhang, 2011), remote sensing (Droogers et al., 2010), and tree growth (Gaucherel et al., 2008). Notable advantages of PEST include that (1) the software can generally complete a parameter estimation process with an extremely high level of model run efficiency (Chen et al., 2016), (2) PEST requires little prior knowledge of programming and (3) PEST can be used on a wide range of mathematical models. Further, the freeware is supplied with a number of utility programs that facilitate iterative parameterisation, e.g. multiple rounds of parameterisation via replacement of optimised parameters into PEST control files (PARREP), addition of parameter prior information (ADDREG), differential weighting of observations (PWTADJ) and several other programs that prevent tedious manipulation of PEST control files by users.

PEST optimises model parameters through successive perturbations in response to the difference between modelled and measured data, within which users may implement local or global optimisers. The default local optimisation scheme uses the Gauss-Levenberg-Marquardt algorithm (Marquardt, 1963), an iterative method that is a hybrid of the Gauss-Newton algorithm and the method of steepest descent. At each step of the iteration, the response surface (ϕ) is approximated by the ϕ value evaluated for the previous parameter set plus the step size multiplied by the Jacobian matrix (J), which is the derivative of the function with respect to the current parameter set. A critical constant implicit to this process is the value of the damping factor λ, which is adjusted from one iteration to the next. If the sum of the squared deviations between observations and measured values S is large, λ is reduced, bringing the algorithm closer to the Gauss-Newton algorithm, whereas if S is small, λ is increased, such that the algorithm approximates the method of gradient descent (Marquardt, 1963). When the rate of convergence is low, as would be the case when the gradient of ϕ approaches zero, λ is increased in response to reduced curvature of the objective function, preventing some of the reduction in parameter increment step size as the algorithm converges on the minimum of the response function. The factor used to adjust λ between successive iterations (RLAMFAC) is one of the variables examined in the present study.

The Gauss-Marquardt-Levenberg (GML) algorithm (hereafter, the 'default') in PEST can be used either with or without Tikhonov regularisation. When properly formulated, mathematically regularised inversion has several advantages, including provision for multiple parameters to be calibrated during the matrix inversion (parameterisation) process. Doherty (2016a) indicates that regularised inversion promulgates minimum error variance, and is numerically stable. It does not founder for want of an invertible matrix as the inverse problem is formulated in a way that guarantees matrix invertibility. Other advantages include the allowance for heterogeneity to emerge in a solution where its existence is supported by data (and suppression of heterogeneity in modelled outcomes where it is not supported by the data), accommodation of model parameter non-uniqueness, and identification of parameter values that cannot be estimated during inversion (singularity) (Doherty 2016a).

Another optimisation algorithm that can be employed using PEST is covariance matrix evolution strategy (CMAES_P). Unlike the default method in PEST, however, CMAES_P does not require derivatives of model outputs with respect to adjustable parameters in order to enable calibration. Thus it can be employed where model outputs show "numerical granularity" due to model numerical solution instability (Doherty 2016a), or where the model is highly nonlinear and/or the response surface shows local minima at various scales (Hansen and Ostermeier, 2001). Evolutionary algorithms are based on the principle of natural selection to guide the evolution towards a global optimum in a discrete or real-valued search space. A population of individuals is created, evaluated

with a pre-defined ϕ, and updated by a combination of operators (selection, recombination, mutation) to create the next generation (Rouchier et al., 2015). This process is repeated until some stopping criterion is met. The principle of CMAES_P is that each generation of ψ individuals is created following a multivariate normal distribution in which the mean and covariance matrices are adapted after the evaluation of the previous generation (Rouchier et al., 2015). After each generation, the mean of the distribution is moved towards previously successful individuals, while the covariance matrix is adapted as to favour previously successful mutation steps in the future. The selection is of type (ω, ψ), in that the ω best individuals of the parent generation determine the creation of a number ψ > ω of offsprings, and no individual from the parent generation is kept unto the next one (Rouchier et al., 2015).

Where model derivatives have integrity, the default gradient-based optimisation processes of PEST are likely to be superior to that of CMAES_P (Doherty 2016a). In contrast, where model derivatives do not have integrity, the performance of CMAES_P may be superior to that of the GML algorithm and/or Tikhonov regularisation (Doherty 2016a). Since this study used a multi-component objective function comprised by several diverse biophysical datasets (e.g. grain yield, soil water content, nitrous oxide emissions etc.), it is likely that the ϕ surface contains discontinuities, local minima, noise, and overall is rugged, such that in several locations of the landscape, model derivatives may not have integrity. If this assumption is true, CMAES_P should result in lower overall ϕ value compared with PEST's gradient-based algorithms.

There are several applications where auto-parameterisation approaches could be used in agricultural modelling scenarios. The first is model-intercomparison studies, such as those documented by Rosenzweig et al. (2013), Lampe et al. (2014) and Ehrhardt et al. (2018). In these studies, users were required to calibrate their model of choice using time-series of measured data that were typically measured in the field (Rosenzweig et al., 2013). However, the extent to which anthropogenic elements and/or user predisposition influenced modelled results in such studies is unknown. Another application of auto-parameterisation is to extensive measured datasets, such as that documented by Field et al. (2016), where manual calibration procedures become too tedious due to the number of measured datasets assumed in the calibration. Use of an automated calibration program such as PEST could potentially remove some of the inherent differences in modelled results caused by differences in user parameterisation techniques and/or knowledge, and could automate standardised numerical recipes for model calibration across diverse datasets such as that described by Field et al. (2016).

Previous studies have shown that PEST can successfully be used for parameterisation of APSIM (Akponikpè et al., 2010; Chen et al., 2016). However, these studies were focussed on APSIM and agronomic results rather than optimisation and applied the default PEST GML algorithm. The extent to which other optimisation methodologies within PEST (Tikhonov regularisation, SVD vs LSQR and CMAES_P) enable calibration of APSIM parameters is yet unknown, as is the CPU time required for optimisation of the multitude of options available to PEST users. The trade-off between optimisation quality and optimisation time is also important. Although optimisation generally improves with computational time, complex evolutionary algorithms such as CMAES_P may require hundreds of hours to run if the number of parameters estimated is large and the model is complex. Continuing optimisation processes for too long may also result in overfitting. On the other hand, premature convergence of an optimisation algorithm may result in poor parameterisation and thus model predictability. Thus, the aim of this study was to determine which PEST algorithms and settings were conducive to the lowest residual difference between APSIM-generated data and measurements in the fewest possible number of model calls. Here our focus was on optimal PEST settings, rather than

optimal APSIM parameters. The purpose of this study was to identify PEST algorithms (via control file settings) that resulted in the best fit of APSIM simulations to measured data through optimisation of APSIM parameters.

Methods

Experimental data

Data were obtained from experiments conducted at Turin, Italy (44° 53'N, 7° 41'E). Replicated measurements were made for nitrous oxide emissions, above-ground biomass, grain yields, cumulative crop nitrogen uptake of above-ground biomass, harvest index, and soil water content; all variables except N_2O were monitored over three years; N_2O was monitored for two years. Four replicated plot measurements were made for each variable except N_2O, which had three to nine replicates per treatment (in the present study, we compared all measured variables to simulated values and fitted APSIM to the means of field measurements). Full details of field experiments are provided in Alluvione et al. (2010), Alluvione et al. (2013) and Grignani et al. (2012); only a brief reprise is given here. The data used for this study were part of a larger experiment with multiple treatments that examined agronomic responses and greenhouse gas emissions of maize crops; here we used the urea treatment detailed in Grignani et al. (2012).

The soil at the experimental site was deep, calcareous, and fertile, and had a silty loam texture. The long-term average yearly temperature is 11.9°C, and the long-term average yearly precipitation is 734 mm. The climate type is F (hot temperate climate without dry season, similar to temperate climates), with two main rainfall periods, in spring and autumn (Supplementary information 1).On the day of sowing each year (19 May 2006, 4 June 2007 and 19 May 2008), experimental plots were prepared by mouldboard plowing at 30 cm deep. Seeds of the FAO 500 maize hybrid PR34N43 (*Zea Mays* L. Pioneer Hi-Bred) were sown 2 cm deep at densities of 7.4 seeds/m^2. Mineral fertiliser as urea at a rate equivalent to 130 kg N/ha was applied at sowing each year. Crops were irrigated throughout the growing season according to evapotranspiration requirements (see Supplementary Information 1). Harvesting was conducted on the 22 September 2009, 10 October 2010 and 29 September 2011. Further details of experimental conditions are provided in Grignani et al. (2012).

Total biomass and N uptake were assessed by hand-harvesting at dent stage from an area of 15 m^2 per plot, with four plot field replicates. Plant samples, separated into grain and shoot/leaves, were oven-dried at 70°C and analysed for N content using a CHN elemental analyser.

Soil NO_3−N content was determined by collecting soil samples before sowing, at flowering, and after harvest from three soil layers (0–15, 15–30, 30–60 cm) in all plots and all years. Soil nitrates were extracted by shaking 100 g of moist soil with 300 mL of 1 M KCl solution for 1 h. Subsequently, the samples were filtered and NO_3 −N concentration was determined by colorimetry with a continuous flow analyser. Soil moisture was measured on the same dates through weighing *c.* 100 g of soil before and after oven drying at 105°C.

Measurements of CO_2, N_2O and CH_4 fluxes were performed through a non-steady state closed chamber technique (Alluvione et al., 2009) coupled with an Innova 1412 photoacoustic infrared gas analyzer (LumaSense Technologies A/S, Ballerup, Denmark). Within each plot three chambers (240 mm; height: 110 mm; wall thickness: 6.2 mm) were monitored for a total of nine measurements per treatment. Fluxes were estimated assuming a linear change of gas concentration over time during

chamber closure and applying proper corrections for fluxes underestimation by the linear model due to the alteration of near-surface concentration gradients (Venterea and Baker, 2008).

Biophysical model for agronomic simulations

APSIM is a biophysical model that simulates the growth and development on a daily time step in response to climate inputs (maximum and minimum daily temperature, solar radiation, rainfall and vapour pressure), soil water, nitrogen, soil organic matter and residue and crop management (Keating et al. 2003). The model is discussed in detail by Keating et al. (2003) and Holzworth et al. (2014). APSIM v7.8 was used to conduct this study. The model was initialised with soil data from Alluvione et al. (2013) (Supplementary information 1); these data included soil water characteristics, organic carbon, pH and soil texture. Crop management conditions in the model were set in line with experimental data described above assuming simulated tillage with discs in the absence of an option to simulate cultivation by mouldboard plows.

As the FAO 500 cultivar used in the field experiments (PR34N43; see Alluvione et al. 2010) was not available in APSIM, a new cultivar was created in the APSIM Maize XML file using the parameters for the "usa_18leaf" variety provided in the default APSIM cultivars (this variety was selected as it had a similar thermal time to maturity as that for FAO 500). APSIM parameter files (located in the C:/Program Files/APSIM directory) for the crop (Maize.xml), soil (Soil.xml) and soil organic matter (SurfaceOrganicMatter.xml) were used to establish the cultivar PR34N43 and associated soil conditions, and later to demark APSIM parameters amendable for modification by PEST (see below). The 'ApsimToSim' executable provided with the default APSIM download package was used to create an APSIM simulation file (.sim) containing all of the APSIM parameters and management information from both the graphical user interface and XML files mentioned above. Availability of the APSIM .sim file containing all of the parameters in each simulation is a key feature allowing APSIM to be optimised by PEST, as is the ability to run APSIM using command prompt arguments specifying the location of the APSIM model executable (ApsimModel.exe in the MS Windows Program Files directory) and the .sim file to be used in each simulation. The ApsimModel executable allows PEST to run APSIM, read model outputs contained in the APSIM .out files, modify specified parameters in the .sim file, rerun APSIM using the modified .sim file, re-evaluate APSIM outputs, and so on. At the start of the parameterisation process, 115 APSIM parameters were identified as having moderate to significant influence on the magnitude of one or more simulation variables; these APSIM parameters were later used as part of the optimisation process (APSIM has much more than 115 parameters). APSIM parameters optimised by PEST were identified through sensitivity analyses wherein each parameter was individually modified by 10% and the magnitude of change in APSIM outputs observed; any APSIM parameter causing more than 10% change in one or more APSIM outputs was used as a basis for selecting a given APSIM parameter for later optimisation in PEST. Optimised APSIM parameters included those influencing the magnitude and thus temporal variability in soil water (e.g. A_to_evap_fact), soil nitrate or ammonium (e.g. solute_flow_eff), phenology (e.g. tt_enjuv_to_init), biomass (e.g. transp_eff_cf) or grain development (e.g. grain_gth_rate). A complete list of APSIM parameters used for optimisation and their bounds are shown in Supplementary Information 2. The same APSIM parameters were optimised for all GML optimisation runs, and a subset of these were analysed for CMAES_P optimisation runs. APSIM parameter bounds were based on author experience with APSIM, as well as APSIM online literature via http://www.apsim.info/. As far as possible, upper and lower parameter bounds were set to physiologically or biophysically meaningful limits, e.g. for the default maize grain growth rate of 9.17 mg/grain.day, we set limits of 0.1 and 10 mg/grain.day. For APSIM parameters expressed as

fractions, such as the root exploration parameter (XF) or the fraction of retained biomass C returned to biomass, we set limits as 1E-9 and 1.0 (PEST does not handle zero value model parameters so instead of zero we set lower bounds to 1E-9). It is important to stress that we chose more parameters than would be chosen in a typical optimisation process. We did this because we had (1) to ensure that all possible sensitive APSIM parameters were included in the optimisation and (2) to determine whether PEST could optimise so many APSIM parameters simultaneously.

Automated parameterisation protocols for multiple objective functions

The freely available model-independent parameter estimation and uncertainty analysis software PEST (http://www.pesthomepage.org/) was used to conduct the automatic parameterisation processes described here. All optimisation runs were conducted by running either PEST, CMAES_P, or other utility programs in the command prompt. Prior to optimisation, PEST requires four main types of files. These include instruction, template, parameter and control files. Instruction files (extension .ins) were created for APSIM outputs corresponding to yield variables (grain yield, final biomass, grain N concentration, total crop N and total grain N), as well as for soil water, N_2O and NO_3 in layers 1-3. Instruction files allow PEST to identify which model outputs correspond to observations, as well as the magnitude to which parameter adjustment influences model outputs. Template files (extension .tpl) were created from APSIM simulation (.sim) files, with hashes (#) for identifying parameters that were amenable for modification by PEST. Parameter files (extension .par) contain initial APSIM parameter values, parameter scaling (1.0 in all cases) and parameter offsets (0.0 in all cases), as well as precision and decimal point notation of PEST computations (the parameter file for treatment 1 in Table 1. Control files (extension .pst; e.g. see Supplementary information 3) contain all of the information required for PEST to be able to run APSIM, read APSIM output files (.out), alter hash-demarked parameters in the APSIM .sim based on APSIM .out files, then repeat the said process based on the changes in parameters and objective function described below. Control files also contain a number of PEST-specific parameters, each present within defined sections. These include PEST convergence criteria, regularisation constraints, measured field data, parameter transformations, upper and lower bounds for APSIM parameters, parameter groups, prior information equations, specification of APSIM output files, etc. In addition to observations (field measurements) and APSIM parameters subject to modification, the PEST control file also contains so-called 'parameter groups' and 'observation groups'. Parameter groups are variables assigned to common APSIM parameters, e.g. we created the parameter group 'kl' that was assigned to APSIM parameters KL1-KL5 (which specify the maximum rate of water extraction in each of the five layers of the soil profile). 'Observation groups' were groups of field variables (13 in total); in this study these included biomass, grain yield, harvest index, nitrous oxide emissions, cumulative crop nitrogen uptake, volumetric soil water content in three layers, grain nitrogen content per unit area, grain nitrogen concentration and soil nitrate concentration in three layers. Initial control files were built using the utility program PESTGEN, while instruction, template and control files were checked for errors using the PEST utility programs INSCHEK, TEMPCHEK and PESTCHEK, respectively (see Doherty, 2016b for details on use of these programs). Many of the PEST parameters/allowable settings in the control file were manipulated to determine the best PEST settings required to obtain the lowest possible objective function (i.e. sum of squared weighted residuals) in the fastest possible computational time. These PEST control parameters and settings are now briefly described, however for a detailed description of each PEST setting in each optimisation run, readers are referred to Doherty (2016a) and Doherty (2016b).

Theory: PEST optimisation algorithms in this study

Three main types of optimisation were employed in this study: the first two included the GML algorithm, the second also included the GML algorithm but also Tikhonov regularisation, and the third included covariance matrix adaptation evolution strategy (PEST implementation abbreviated as CMAES_P). The GML algorithm is a gradient-based optimisation approach, whereas CMAES_P is a "genetic-type" algorithm that does not employ derivatives to conduct optimisation.

Each of the three main optimisation types were tested with and without one or more minor forms of regularisation. For the purpose of this study, regularisation is a means through which a unique solution is obtained to an inverse problem where the calibration dataset lacks the information to support uniqueness (Doherty, 2016a), i.e. the situation wherein only one combination of model parameters provides the lowest difference between measurements and modelled values. For each of the three main optimisation processes, we examined the effect of adding either singular value decomposition (SVD) or least squares with QR decomposition (LSQR). In contrast to Tikhonov regularisation, which is a major form of regularisation and enables solution of an ill-posed problem by adding information derived from initial parameter estimates (prior information), SVD and LSQR are minor forms of regularisation that remove model parameter combinations from the problem by subdividing the estimated model parameters into two orthogonal subspaces, one comprising the "calibration solution subspace" and the other the "calibration null space", the latter of which is spanned by model parameter combinations that cannot be estimated during an inversion process (C4SF, 2017). In PEST, Tikhonov regularisation can be applied in conjunction with either GML or CMAES_P optimisation and either without or without SVD or LSQR (but SVD and LSQR cannot be combined in any given optimisation).

In this section, we first briefly describe the theoretical background of each of the three optimisation algorithms, then discuss further background to SVD and LSQR. Both optimisation and regularisation algorithms are presented in the context of implementation within the PEST framework.

1. Gauss-Marquardt-Levenberg algorithm (the default)

The GML algorithm used for the default optimisation runs ("estimation mode" in PEST) computes an objective function (ϕ_m) based on nonlinear least-squares minimisation between the response surface from the model and the measured data (the 'm' subscript denotes measured data). The GML is a gradient-based approach, and as such, may only find local minima. Model parameters are calculated in an iterative fashion as PEST systematically varies model inputs, runs the model, reads the model output, and evaluates the model fit using ϕ_m, which represents the weighted least squares difference between observed and simulated values (Doherty and Hunt, 2010). The objective function for the GML algorithm can be expressed as:

$$\phi_m = [\mathbf{c} - \mathbf{Xa}]^T \mathbf{Q_m} [\mathbf{c} - \mathbf{Xa}] \quad \textbf{(1)}$$

where $\mathbf{Q_m}$ is a diagonal matrix whose i^{th} element q_{ii} is the square of the weight w_i attached to the i^{th} field measurement, **c** is a vector of measured values, **a** is a vector of APSIM parameters to be estimated, **X** is a matrix of APSIM outputs based on parameter vector **a** and collocated with the observations in **c**, and ***T*** indicates matrix transpose. Following Lin (2005), the parameter vector **a** is updated on iteration $j + 1$ using:

$$\mathbf{a}^{j+1} = \mathbf{a}^j + \rho.[\mathbf{J}^T\mathbf{J} + \lambda\mathbf{B}]^{-1} \times \mathbf{J}^T(\mathbf{c} - \mathbf{Xa}) \quad \textbf{(2)}$$

where $\mathbf{a}^j$ is a vector of estimated APSIM parameters on the j^{th} iteration, **J** represents the Jacobian (matrix containing all first-order partial derivatives of simulated values that correspond to

observations in the calibration dataset to the adjustable model parameters $\mathbf{a}^j$), ρ is a PEST parameter between 0 and 1 which is chosen so that $\phi_m(\mathbf{a}^{j+1}) < \phi_m(\mathbf{a}^j)$, **B** is a diagonal matrix with elements taken from $\mathbf{J}^T\mathbf{J}$, and λ (the Marquardt lambda) is computed numerically during each iteration.

Equation (1) can be alternatively expressed as

$$\phi_m = \sum_{i=1}^{n} [w_i r_i]^2 \quad \textbf{(3)}$$

Where r_i (the i^{th} residual) is the difference between the modelled and measured value for the i^{th} measured variable and w_i is the corresponding weight matrix attributed to the i^{th} residual. n represents the total number of observation groups. Thus, in this study, the 13 components of ϕ_m included grain yield, biomass, harvest index, grain N concentration, cumulative crop N uptake, grain N content, volumetric soil water content in three layers, soil nitrate concentration in three layers and soil nitrous oxide emissions. For treatments 1-24, each weight w_i was assigned using the PEST utility program PWTADJ2 such that weights were inversely proportional to the standard deviation of each 'observation group' in the PEST control file (there were 13 observation groups). Weights were uniformly assigned within observation groups but differentially across observation groups. Weighting in this way defends the inversion process against one or more observation groups with high standard deviation dominating the value of ϕ_m. Weighting applied in all treatments is shown in Supplementary information 5.

At the start of each iteration, the relationship between the best model parameters and model outputs is linearised using a Taylor-series expansion. The finite-difference method is used to compute the Jacobian matrix (Necpálová et al., 2015). The linearised solution is then solved for the updated model parameter set using the GML algorithm, and the new ϕ_m is calculated as defined above. The model parameter changes and value of ϕ_m are compared with those of the previous iteration to determine if another iteration is justified. If it is, the entire process is repeated; if not, the parameter estimation process terminates (Doherty, 2016a).

In PEST, the real variable RLAMBDA1 is the initial value of λ (Eqn. 2). In general, the value of λ should decrease as the number of iterations increases. The effect of RLAMBDA1 was tested because the initial value may have an impact on the rate of convergence of the algorithm and the final value of the objective function. Doherty (2016a) indicates that ill-posed problems are more likely to result in singularity in matrix inversion (singularity prevents matrix inversion and thus derivation of optimal parameter vectors). For such problems, increasing the value of RLAMBDA1 to 10 (from the default of 5) and setting the value of RLAMFAC to -3 (the factor by which PEST adjusts λ as it tests different values of this variable for their efficacy in lowering ϕ). By setting RLAMFAC to -3, PEST adjusts λ during each iteration of the inversion process so that λ can achieve a value of 1.0 with three adjustments. This allows rapid adjustment of λ if local parameter insensitivity promulgates sudden problem ill-posedness (Doherty, 2016a).

2. <u>Gauss-Marquardt-Levenberg algorithm with Tikhonov regularisation</u>

The second main optimisation algorithm employed in this study also used the GML algorithm, but included Tikhonov regularisation (treatments 25-54). Mathematical "regularisation" is the process of adding information into an optimisation search to solve an ill-posed problem and to prevent over-fitting. To conduct optimisation runs using Tikhonov regularisation, PEST must be run in "regularisation" mode, wherein PEST defines two objective functions instead of only one defined in "estimation" mode (Eqn. 1). The objective function in "regularisation" mode is comprised by the

measurement objective function, designated ϕ_m, and the regularisation objective function, designated ϕ_r. This constitutes a weighted least-squares measure of the discrepancies between the model parameters and their preferred conditions:

$$\phi_r = [\mathbf{d} - \mathbf{Za}]^T \mathbf{Q}_r[\mathbf{d} - \mathbf{Za}] \quad \textbf{(4)}$$

where ϕ_r is a diagonal matrix of the squares of weights assigned to the various "regularisation observations" which comprise vector **d**. The relationships between the regularisation observations in **d** and their model-generated counterparts (calculated from model parameter vector **a**) are encapsulated in matrix **Z** (Doherty, 2016a).

To assign every APSIM parameter with a preferred value equal to its initial value, the ADDREG1 utility program described in Doherty (2016b) was used. The ADDREG1 program adds a series of prior information equations to the PEST control file that are assigned to PEST parameter groups beginning with "regul_" in the "prior information section" (see Supplementary information 3). Collectively, the addition of prior information equations using ADDREG1 comprises a Tikhonov regularisation scheme (Doherty, 2016b). In essence, prior information equations constitute a set of observations which pertain directly to the model parameters themselves. As such, they comprise part of the calibration dataset which assists in the estimation of APSIM parameters. Using ADDREG1, one linear prior equation is added for each APSIM parameter cited in the control file. In each prior equation, the APSIM parameter is set equal to its initial value (or the log of its initial value if the APSIM parameter is transformed). Similar to individual observations in the PEST control file, weights must be assigned to each prior equation; these weights are multiplied internally by a regularisation weight factor (μ) before formulation of an overall ϕ during each iteration of the inversion process (Eqn. 6 below). Treatments 25-54 had 115 prior information equations and 69 observation groups (56 of which were associated with prior information; e.g. see Supplementary information 3). All prior information equations were assigned a weight of 1.0 (the default). Setting the PEST variable IREGADJ to 1.0 allows PEST to vary the regularisation weights between groups, thus complementing the information density of the calibration dataset (Doherty, 2016b).

By way of example of prior information, for the APSIM soil nitrogen denitrification parameter we created the PEST parameter 'dnitcof' and a corresponding regularisation group parameter called 'regul_dnitco' (all APSIM parameters optimised by PEST must be represented by corresponding PEST parameter names less than or equal to 12 characters in length). The prior information equation thus created using ADDREG1 was:

$$\log_{10}(\text{dnitcof}) = -2.97 \quad \textbf{(5)}$$

where the log was introduced as the PEST parameter denitcof for this example was log transformed prior to the inversion process (see Supplementary information 3) and -2.97 represents the log of the initial dnitcof value (1.05E-03). Throughout the optimisation process, the extent to which dnitcof differs from 1.05E-3 causes a non-zero residual, and the value of ϕ_r in Eqn. 5 becomes non-zero.

To prevent over-fitting, the user is required to provide a target measurement objective function (ϕ^t_m). This is the value of PHIMLIM shown in Tables 1-4 (set to 1.00E-10 in treatments 25-54). PEST attempts to minimise the value of ϕ_r subject to the constraint provided by ϕ^t_m. In solving this constrained minimisation problem, PEST applies a global multiplier to all weights that are ascribed to prior information equations (Doherty 2016a). During each iteration of the inversion process, PEST minimises the total objective function:

$$\phi = \phi_m + \mu^2\phi_r \quad \textbf{(6)}$$

where μ is the regularisation weighting factor. During each iteration, PEST computes the optimal value of μ. Under the linearity assumption used to compute the Jacobian matrix, this is the value of μ that results in a model parameter upgrade vector for which ϕ_m is reduced to a value as close as possible to ϕ_m^t. When PEST is not able to lower ϕ_m to ϕ_m^t, it accepts the upgraded model parameters and proceeds to the next iteration. However, if PEST does succeed in lowering ϕ_m to an acceptable level, it then attempts to lower ϕ_r while maintaining ϕ_m below this acceptable level. This acceptable level is the variable PHIMACCEPT and should be set slightly higher than ϕ_m^t (the default PHIMACCEPT value is 1.05E-10).

The PEST parameter FRACPHIM shown in Tables 1-4 represents the new value for ϕ^t_m calculated at the beginning of every iteration; this value is calculated as the current value of ϕ_m times FRACPHIM, or the current value of FRACPHIM, whichever is greater. FRACPHIM was set to 0.1 for treatment 25. WFINIT, WFMIN and WFMAX are the initial, minimum and maximum permissible regularisation weight factors, respectively. PEST parameter WFFAC defines the multiplier used to adjust the regularisation weight factor such that the value of ϕ_m equals that of ϕ_m^l, whilst PEST parameter WFTOL defines the maximum allowed difference between two successive weighting factors (Doherty 2016a). The variable IREGADJ is used to adjust the weighting factor within regulation groups. When it is set to 1, PEST multiplies the weights pertaining to all members of each regularisation group by a group-specific factor. This factor is chosen so that the total composite sensitivities of all regularisation groups are the same. It is important to note, however, that relative weighting within each observation group remains unchanged when IREGADJ equals 1.0 (Doherty 2016a).

3. Covariance matrix adaptation-evolution strategy

The PEST control files for the GML algorithms were also compatible for use with the third main algorithm employed in this study: covariance matrix adaptation-evolution strategy (PEST implementation abbreviated CMAES_P). In contrast to the GML algorithm, CMAES_P does not apply gradient-based methods, and thus is theoretically capable of finding the global minimum of the search space. In CMAES_P, a population of new search points ($\psi \geq 2$) is generated by sampling a multivariate normal distribution. The basic equation for sampling the search points for generation number g reads:

$$\boldsymbol{x}_k^{(g+1)} \sim \boldsymbol{m}^{(g)} + \boldsymbol{\sigma}^{(g)}\boldsymbol{v}(\mathbf{0}, \boldsymbol{C}^{(g)}) \qquad \text{for } k = 1, \ldots, \psi \qquad \textbf{(7)}$$

where ~ denotes the same distribution on the right and left hand sides, $\boldsymbol{v}(\mathbf{0}, \boldsymbol{C}^{(g)})$ is a multivariate normal distribution with zero mean and covariance matrix of the search distribution $\boldsymbol{C}^{(g)}$, $\boldsymbol{x}_k^{(g+1)}$ represents the k^{th} offspring (individuals, search points) from generation $g + 1$, $\boldsymbol{m}^{(g)}$ represents the weighted average value of the search distribution of ω selected parents ($\omega < \psi$) at generation g, and $\boldsymbol{\sigma}^{(g)}$ is the overall standard deviation (step-size). The number of generations g depends on CMAES_P termination criteria that are prescribed by the user. Further details of CMAES and supporting theoretical background are described in Hansen (2016).

For simplicity, several CMAES_P optimisation criteria were not altered from their defaults (Table 4). These included the random number seed (1111) for initialisation, the minimum relative objective function or model parameter change over 40 iterations (1.00×10^{-3}), the relative high-low objective function difference over 10 iterations (1.00×10^{-2}), and the maximum iteration count (1000; see Table 4). For CMAES_P, we examined the effect of ψ, ω, singular value thresholds, singular value decomposition (SVD) hybridisation, log transformation of model parameters, inclusion of prior

information, and whether or not model parameters causing model run failures should be weighted lower than other model parameter vectors.

For CMAES_P, weights corresponding to ω model parameter sets can be assigned as “super-linear”, “linear”, or “equal”. In the first two cases, greater weight is given to model parameters that give rise to lower ϕ values, this often leading to faster reduction of the objective function. Since the first two cases tend to elicit similar responses in model optimisation runs, this study only examined “superlinear” and “equal” weighting (the latter in treatment 60). Once a new average set of model parameter values has been computed in this fashion, the next iteration begins. Because random model parameter realisations are generated to be symmetrical about this mean, there is a tendency for the objective function to fall as iterations proceed (Doherty 2016a). A caveat of CMAES_P is run time burden in optimisation runs that include multiple model parameters. In this study, CMAES_P would not allow simultaneous optimisation of 115 APSIM parameters, so the number of optimised parameters was reduced from 115 in GML optimisation runs (treatments 1-54), to 84 in treatments performed by CMAES_P (55-71). Accordingly, the number of PEST parameter groups in CMAES_P treatments was reduced to 39. The 31 APSIM parameters removed from the GML control files in preparation for the CMAES_P runs were chosen according to their sensitivity. These APSIM parameters were identified from the PEST “.sen” files that were produced after each GML optimisation run (APSIM parameter sensitivity was consistent regardless of treatment applied). Inspection of .sen files showed that insensitive APSIM parameters were not modified by PEST during GML or Tikhonov regularisation runs. Thus, it is likely that the fewer parameters contained in the CMAES_P treatments had little effect on the final degree of fit achieved. APSIM parameters optimised using CMAES_P are shown in Supplementary Information 2.

At the CMAES_P prompt, users must select whether “soft” or “hard” hybridisation takes place. “Soft” hybridisation replaces the best of the currently-selected ψ parameters (these forming part of the m + 1 member parameter set on which SVD analysis was based) with the SVD-computed parameter set if the value of ϕ achieved through SVD yields the lowest ϕ to date. If the “hard” option is selected, parameter set replacement is undertaken if the SVD-computed parameter set leads to a lower ϕ than that computed only on the basis of the current ψ parameter sets (Doherty 2016a).

4. Minor regularisation methods: singular value decomposition (SVD) and least-squares with QR decomposition (LSQR)

In PEST either singular value decomposition (SVD) or LSQR (Least Squares with QR decomposition) can be combined with any other optimisation algorithm. Both methods were originally developed for the inversion of ill-conditioned matrices (Lanczos, 1961; Paige and Saunders, 1982). In contrast to the analytical approach afforded by SVD, LSQR is an iterative numerical approach designed for inversion of large matrices. Although LSQR generally allows faster convergence, it is an approximate measure and thus may not result in ϕ values that are as low as those obtained using SVD. Hence, we investigated the influence of both SVD and LSQR on the value of ϕ and computational time for each of the three main algorithms described above.

SVD is a form of matrix factorisation into rotational and scaling matrices, enabling tractability to the solution of ordinary least-squares problems in matrix inversion by preventing matrix singularity. Truncated SVD of the weighted Jacobian matrix in PEST occurs on an iteration-by-iteration basis (Necpálová et al., 2015). The level of truncation was automatically calculated based

on a stability criterion. SVD transforms the original model parameters into linear combinations (i.e., eigenvectors), determines which are most sensitive (James and John, 2005; Moore and Doherty, 2006), and truncates the transformed normal equations matrix, reducing the number of estimated parameters to maintain numerical stability and maximum reasonableness (Aster et al., 2013). The resulting regularised inversion process will not include parameters that are unidentifiable with the available data. When correlated parameters are included in the inversion, the SVD-based regression finds the maximum likelihood combination of the parameters that is consistent with the observations (Necpálová et al., 2015). In all SVD treatments in this study, the PEST variable SVDMODE was set to 2, such that PEST undertook singular value decomposition of the $\mathbf{Q}^{1/2}\mathbf{J}$ matrix, where **Q** is a weighting matrix and **J** represents the Jacobian matrix described above.

The LSQR algorithm (Least Squares with QR decomposition) represents another mechanism that can be used to solve inverse problems (Paige and Saunders, 1982). LSQR attempts to subdivide parameter space into orthogonal null and solution spaces, and then restricts solution of the inverse problem to the latter space (Paige and Saunders, 1982). Because LSQR facilitates matrix sparsity and compartmentalisation of the solution into a matrix subspace (rather than attempting to linearise the entire solution as conducted by SVD), LSQR tends to converge much faster than SVD (Lin et al., 2016).

Treatments conducted in this study

Seventy-one treatments were conducted. These examined various PEST control file settings for each of the optimisation algorithms and regularisation techniques presented above. Table 1 presents a brief description of the PEST parameters examined in this study, while Tables 2-4 show the values of PEST parameters compared with their default values. Treatments 1-24 describe PEST optimisation runs conducted in the 'estimation' mode (using the GML theorem), Treatments 24-54 were conducted in 'regularisation' mode and thus used the GML algorithm with Tikhonov regularisation, while treatments 55-71 were conducted using CMAES_P (Table 4). Descriptions in Table 1 provide a minimal level of background required to enable understanding of the concepts used in this study; more detail regarding PEST parameters and theory underlying each treatment is shown in Supplementary information 4, Doherty (2016a) and Doherty (2016b).

Treatment 0 contained only management and soil information measured at the site; no parameterisation was conducted for this treatment. PEST was used to compute ϕ for this treatment using the control file from treatment 1 (see below) but with the number of optimisation runs set to zero (using PEST parameter NOPTMAX in the control file). This treatment was not judged as the baseline because the maize hybrid used in the field trials was not available in APSIM; as such, part of the calibration process in all treatments involved optimising parameters for the new hybrid in APSIM.

The PEST control file for treatment 1 was used to create the default treatment (baseline) upon which all other treatments were compared. Initial weights applied to observations in this file and in all subsequent files were established using the utility program PWTADJ2 according to observation groups, wherein weights were assigned to each observation group that were the inverse of the standard deviation associated with the corresponding observation group. Group specific calibration weight adjustment was employed using PWTADJ2 to accommodate the fact the PEST would likely experience more difficulties in fitting some modelled values to field measurements compared with

others. Weighting applied by PWTADJ2 was retained for all treatments except treatments 52 and 53, which were designed to examine the effect of observation group weighting. Treatment 52 adopted the optimised parameters from one of the better performing treatments (treatment 43), then increased weighting applied to the N_2O group, since results showed that relative contribution to ϕ from this group was large. The weighting applied to all N_2O observations was increased from 4.203 (as for previous treatments) to 20, with the rationale that PEST would thus "focus" on reducing the error between modelled data and measurements of this observation group. Similar to treatment 52, treatment 53 reused the optimised parameters resulting from treatment 43. The PWTADJ1 utility program provided with PEST was used to re-adjust the weighting applied to all datasets, such that the total contribution of all datasets to ϕ was 10. A summary of observation group weighting for all treatments is provided in Supplementary information 5.

For treatments 1-54, 115 APSIM parameters were demarked within the PEST template file by hashes (#) along with a user-assigned parameter name. Corresponding upper and lower bounds were specified for each of these parameters in the PEST control file and were not altered between treatments. For all treatments in Tables 2 and 3, there were 117 field measurements, 56 parameter groups, 0 prior equations, and 13 observation groups in the PEST control files.

To identify sensitive PEST parameters in the control file, groups of only two or three PEST parameters were modified from the baseline file on a piecemeal basis to test the effect of alternative setting groups on the value of the objective function and total run time. However, some of the parameters in the PEST control file required more than one PEST parameter to be modified (e.g. in the 'parameter groups' section, the use of split derivatives required three settings to be simultaneously modified from the baseline file). After key PEST control file parameters causing a significant effect on optimisation time or the objective function deviation from the default value (or both) were identified, combinations of up to five PEST parameters in the control file were modified and tested to determine whether the combined effect of sensitive PEST parameters on total model calls and objective function value was additive or otherwise. Thus, treatments that employed Tikhonov regularisation (Table 3) were constructed based on previous runs without regularisation that reduced either ϕ or the total number of model calls. It should also be noted that although this study explores and extensive number of PEST parameter combinations, not all possible parameter combinations were explored.

Model evaluation criteria

The quality of fit and time for optimisation convergence of each treatment was evaluated using ϕ, CPU time and Pearson's correlation coefficient (r). Both ϕ and r include data from all 13 APSIM variables fitted in each treatment; both variables were computed by PEST.

Table 1 Treatments used to examine PEST control file parameters (table layout follows PEST control file). Parameter descriptions are summarised from Doherty et al (2016a). Variables in square brackets are optional in PEST.

PEST control file parameter	Treatment No.	Type	Values	Description
Control data				
RLAMBDA1	2	Real	≥ 0	Initial λ value at start of optimisation
RLAMFAC	2	Real	positive or negative, but not zero	Dictates λ adjustment process from one iteration to the next
PHIRATSUF	2	Real	0-1	Fractional objective function sufficient for end of current iteration
PHIREDLAM	23	Real	0-1	Termination criteria for λ search
NUMLAM	23	integer	≥ 1	Maximum number of λ values to be tested
[JACUPDATE]	5, 49-50	integer	≥ 0	Activation of Broyden's Jacobian update procedure (mechanism for improving Jacobian matrix)
[LAMFORGIVE]	5, 49-50	Text	"lamforgive" or "nolamforgive"	Assign a high objective function value to any λ search resulting in an APSIM run failure
[DERFORGIVE]	5, 49-50	Text	"derforgive" or "noderforgive"	Accommodates model failure whilst computing the Jacobian by setting pertinent parameter sensitivities to zero
RELPARMAX	3	Real	> 0	Stipulates maximum relative parameter change limit from one iteration to the next
FACPARMAX	3	Real	> 1	Stipulates maximum factor-based parameter change from one iteration to the next
FACORIG	24	Real	0-1	Imposes a minimum factor-based change on parameters that are very small, ensuring sufficient parameter perturbation during inversion
PHIREDSWH	3	Real	0-1	Sets a value for the relative change in objective function for one iteration to the next that stipulates introduction of 3- or 5-point derivatives (cf. forward-differencing derivative default)
[DOAUI]	7, 8	Text	"aui", "auid" or "noaui"	Implements automatic user intervention (AUI; mechanism for defending inversion process against poor finite-difference derivates)
[DOSENREUSE]	6-8, 20, 70	Text	"senreuse" or "nosenreuse"	Reuse parameter sensitivities as opposed to the default of recalculating for each optimisation iteration
[BOUNDSCALE]	9	Text	"boundscale" or "noboundscale"	If SVD is activated, "boundscale" scales all parameters by their upper and lower bounds prior to inversion
NOPTMAX	22-23	integer	-2, -1 or > 0	Number of optimisation iterations
PHIREDSTP	4, 22-23	Real	> 0	Minimum relative change in objective function for triggering termination of optimisation
NPHISTP	4, 22-23	integer	> 0	Number of successive iterations over which PHIREDSTP applies
NPHINORED	4, 22-23	integer	> 0	Number of iterations since last reduction in objective function to trigger termination
RELPARSTP	4, 22-23	Real	> 0	Minimum relative parameter change triggering termination of the optimisation process
NRELPAR	4, 22-23	integer	> 0	Number of successive iterations over which RELPARSTP applies

Sensitivity reuse				
SENRELTHRESH	6-8, 20, 38-39, 41, 44, 46, 48-49, 70	Real	0-1	Relative parameter sensitivity below which sensitivity reuse is activated for a parameter
SENMAXREUSE	6-8, 20, 38-39, 41, 44, 46, 48-49, 70	integer	≥ 1	Maximum number of reused sensitivities per iteration
SENALLCALCINT	6-8, 20, 38-39, 41, 44, 46, 48-49, 70	integer	> 1	Iteration interval at which all sensitivities are recalculated
SENPREDWEIGHT	6-8, 20, 38-39, 41, 44, 46, 48-49, 70	Real	any number	Weight to assign to prediction in computation of composite parameter sensitivities to determine sensitivity reuse
SENPIEXCLUDE	6-8, 20, 38-39, 41, 44, 46, 48-49, 70	Text	"yes" or "no"	Include/exclude prior information when computing composite parameter sensitivities to determine sensitivity reuse
Singular value decomposition				
SVDMODE	9-10, 35, 39, 40, 45-46, 49-50, 67, 70	integer	0 or 1	If SVDMODE is set to 1, activates truncated SVD for solution of inverse problem
MAXSING	9-10, 35, 39, 40, 45-46, 49-50, 67, 70	integer	> 0	Number of singular values before truncation
EIGTHRESH	9-10, 35, 39, 40, 45-46, 49-50, 67, 70	Real	≥ 0 and < 1	Ratio of the lowest to the highest eigenvalue of the (**J**t**QJ** + **λI**) matrix at which singular value truncation occurs (see text in Methods)
EIGWRITE	9-10, 35, 39, 40, 45-46, 49-50, 67, 70	integer	0 or 1	Determines whether SVD file resulting from PEST inversion process is written to text file
LSQR				
LSQRMODE	11, 20, 36, 41-44, 47-48, 52-54, 68-69	integer	0 or 1	Activates LSQR solution of the inversion problem
LSQR_ATOL	11, 20, 36, 41-44, 47-48, 52-54, 68-69	Real	≥ 0	Estimate of the relative error in the data defining the **Q1/2J** matrix used in LSQR (see text in Methods)
LSQR_BTOL	11, 20, 36, 41-44, 47-48, 52-54, 68-69	Real	≥ 0	Estimate of the relative error in the data defining the parameter vector **a** in Eqn. 1
LSQR_CONLIM	11, 20, 36, 41-44, 47-48, 52-54, 68-69	Real	≥ 0	Upper limit of the matrix condition number during the inversion process (higher condition numbers indicate ill-posedness)
LSQR_ITNLIM	11, 20, 36, 41-44, 47-48, 52-54, 68-69	integer	> 0	Upper limit of the number of iterations permitted when LSQR is employed
LSQRWRITE	11, 20, 36, 41-44, 47-48, 52-54, 68-69	integer	0 or 1	Writes output from the LSQR solver to an output file
Automatic user intervention				
MAXAUI	7-8	integer	≥ 0	Maximum number of automatic user interventions per optimisation iteration
AUISTARTOPT	7-8	integer	≥ 1	Optimisation iteration at which to commence automatic user intervention
NOAUIPHIRAT	7-8	Real	0-1	Relative objective function reduction threshold triggering automatic user intervention
AUIRESTITN	7-8	integer	≥ 0 (≠ 1)	Automatic user intervention pause interval expressed in optimisation iterations

AUISENSRAT	7-8	Real	> 1	Composite parameter sensitivity ratio triggering automatic user intervention
AUIHOLDMAXCHG	7-8	integer	0 or 1	When implemented, instructs PEST to hold specific parameters based on their relative change during previous optimisation iterations
AUINUMFREE	7-8	integer	> 0	Cease automatic user intervention if the number of adjustable parameters has been reduced to AUINUMFREE
AUIPHIRATSUF	7-8	Real	0-1	Ratio of objective function computed using AUI to that computed without AUI. If AUIRATSUF is less than this value, implementation of automatic user intervention is terminated
AUIPHIRATACCEPT	7-8	Real	0-1	Relative objective function reduction threshold for acceptance of automatic-user intervention-calculated parameters
NAUINOACCEPT	7-8	integer	> 0	Number of iterations since accepting previous parameter change that triggers termination of automatic user intervention
Parameter groups				
INCTYP	12	Text	"relative", "absolute", "rel_to_max"	Method by which parameter increments are calculated
DERINC	13	Real	> 0	Absolute or relative parameter increment (added or multiplied to existing parameters depending on the value of INCTYP)
DERINCLB	14	Real	≥ 0	Absolute lower bound of relative parameter increment
FORCEN	15	Text	"switch", "always_5"	Determines when higher order derivatives are undertaken for each parameter group (always_5 = five-point derivatives are used, switch = start by using forward difference derivatives then switch to three-point derivatives for all parameter group members on the first occasion that the relative reduction in the objective function between iterations is less than the value of PHIREDSWH)
DERINCMUL	16	Real	> 0	Derivative increment multiplier when undertaking derivatives using methods other than the default forward-differencing method
DERMTHD	15, 17	Text	"parabolic", "minvar" or "best_fit"	Method used to calculate derivatives ("min_var" means minimum error variance; must be implemented with "always_5", "best_fit" is a regression approach implemented with "switch")
[SPLITTHRESH]	18, 49-50	Real	> 0 (0 = deactivation of split slope analysis)	Slope threshold for split slope analysis (an option for mitigating effects of poor model numerical performance on PEST performance wherein segmented analysis is used to compute the change in each parameter)
[SPLITRELDIFF]	18, 49-50	Real	> 0	Relative slope difference threshold allowing implementation of split slope analysis
[SPLITACTION]	18, 49-50	Text	"smaller"	The slope segment with higher absolute value is rejected, and the derivative is taken as the slope of the segment with lesser absolute slope
Parameter data				
PARTRANS	19, 37-46, 49, 52-54, 56-70, 72	Text	"log" or "none"	Parameter transformation prior to inversion ("log" = log to the base 10)
PARCHGLIM	19, 21, 37-46, 49, 52-54, 56-70, 72	Text	"relative" or "factor"	Determines whether optimised model parameters are adjusted relatively or by multiplication of a factor

Regularisation				
PHIMLIM	66-71	Real	> 0	Target measurement objective function (see text in Methods)
PHIMACCEPT	26, 66-71	Real	> PHIMLIM	Acceptable measurement objective function (see text in Methods)
[FRACPHIM]	27, 66-71	Real	≥ 0 (< 1)	Sets target measurement objective function at this fraction of current measurement objective function
[MEMSAVE]	28	Text	"memsave" or "nomemsave"	Activates conservation of memory at cost of execution speed and quantity of model output
WFINIT	29, 66-71	Real	> 0	Initial regularisation weight factor (see text in Methods)
WFMIN	66-71	Real	> 0	Minimum regularisation weight factor
WFMAX	66-71	Real	> WFMIN	Maximum regularisation weight factor
[LINREG]	51	Text	"linreg" or "nolinreg"	Instructs PEST that regularisation constraints are linear or nonlinear, respectively
[REGCONTINUE]	32	Text	"continue" or "nocontinue"	Instructs PEST to continue minimising regularisation objective function even if measurement objective function is less than PHIMLIM (see text in Methods)
WFFAC	30, 66-71	Real	> 1	Regularisation weight factor adjustment (see text in methods)
WFTOL	31, 66-71	Real	> 0	Convergence criterion for regularisation weight factor calculated during each iteration
IREGADJ	33, 34, 66-71	integer	1, 2 or 4	Instructs PEST to perform inter-regularisation group weight factor adjustment, or to compute new relative weights for regularisation observations and prior information equations (see text in Methods)
[NOPTREGADJ]	34	integer	≥ 1	The number of consecutive optimisation iterations stipulating recalculation of regularisation weight factor
[REGWEIGHTRAT]	34	Real	≥ 1	The ratio of the highest to lowest regularisation weight (see text in Methods)
CMAES_P				
Population size (ψ)	61, 65, 69-72	integer	[4 + 3 × ln(n)]	Number of random realisations of n-dimensional parameter vectors generated during each iteration of CMAES_P (n = the number of parameters being estimated)
Number of parents (ω)	62-63	integer	ψ/2	Number of objective function values used to calculate m in Eqn. 7 for the next iteration. Default value is half the population size.
Recombination weights	60	Text	"linear", "superlinear" or "equal"	Weighting given to the lowest objective function values in forming m (Eqn. 7) for the next iteration
SVD-hybridisation	57-59	Text	"soft" or "hard"	Uses all or a subset of the current iteration ψ parameter sets to compute approximate SVDs (see text in Methods)
No. singular value trial thresholds	57-59	integer	≥ 1	Determines level of single value truncation if SVD-hybridisation is employed
Forgive model run failure	64-65, 69-70, 72	Text	"yes" or "no"	'Yes' allows CMAES_P to continue if any parameter set causes an APSIM run failure, "no" ceases CMAES_P if a given parameter set causes an APSIM run failure

Table 2 Data from PEST control files used to conduct optimisation with the GML algorithm. All optimisation runs were performed using 'estimation' mode in PEST, with 115 parameters, 56 parameter groups, 13 observation groups and no prior equations. Parameters shown in bold indicate variation from the baseline file (treatment 1). PEST control settings in the first column are described in the methods, and parameters in square brackets [] indicate optional use in the PEST control file. Each control file section is identified with an asterisk. NA = Not Applicable.

Treatment Nº.	**Baseline**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
* control data											
RLAMBDA1	5	**10**	5	5	5	5	5	5	5	5	5
RLAMFAC	2	**-3**	2	2	2	2	2	2	2	2	2
PHIRATSUF	0.3	**0.1**	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
PHIREDLAM	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
NUMLAM	10	10	10	10	10	10	10	10	10	10	10
[JACUPDATE]	NA	NA	NA	NA	**999**	NA	NA	NA	NA	NA	NA
[LAMFORGIVE]	NA	NA	NA	NA	**lamforgive**	NA	NA	NA	NA	NA	NA
[DERFORGIVE]	NA	NA	NA	NA	**derforgive**	NA	NA	NA	NA	NA	NA
RELPARMAX	3	3	**2**	3	3	3	3	3	3	3	3
FACPARMAX	3	3	**2**	3	3	3	3	3	3	3	3
FACORIG	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
PHIREDSWH	0.1	0.1	**0.2**	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
[DOAUI]	NA	NA	NA	NA	NA	NA	**auid**	**aui**	NA	NA	NA
[DOSENREUSE]	NA	NA	NA	NA	NA	**senreuse**	**senreuse**	**senreuse**	NA	NA	NA
[BOUNDSCALE]	NA	NA	NA	NA	NA	NA	NA	NA	**boundscale**	NA	NA
NOPTMAX	50	50	50	50	50	50	50	50	50	50	50
PHIREDSTP	0.01	0.01	0.01	**0.005**	0.01	0.01	0.01	0.01	0.01	0.01	0.01
NPHISTP	3	3	3	**5**	3	3	3	3	3	3	3
NPHINORED	3	3	3	**5**	3	3	3	3	3	3	3
RELPARSTP	0.01	0.01	0.01	**0.005**	0.01	0.01	0.01	0.01	0.01	0.01	0.01
NRELPAR	3	3	3	**5**	3	3	3	3	3	3	3
* Sensitivity reuse											
SENRELTHRESH	NA	NA	NA	NA	NA	**0.15**	**0.15**	**0.15**	NA	NA	NA
SENMAXREUSE	NA	NA	NA	NA	NA	**-1**	**-1**	**-1**	NA	NA	NA
SENALLCALCINT	NA	NA	NA	NA	NA	**3**	**3**	**3**	NA	NA	NA
SENPREDWEIGHT	NA	NA	NA	NA	NA	**-1**	**-1**	**-1**	NA	NA	NA
SENPIEXCLUDE	NA	NA	NA	NA	NA	**Yes**	**yes**	**yes**	NA	NA	NA
* Singular value decomposition											
SVDMODE	NA	NA	NA	NA	NA	NA	NA	NA	**2**	**2**	NA

MAXSING	NA	NA	NA	NA	NA	NA	NA	NA	**115**	**115**	NA
EIGTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	**5E-07**	**5E-07**	NA
EIGWRITE	NA	NA	NA	NA	NA	NA	NA	NA	**0**	**0**	NA
* lsqr											
LSQRMODE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1**
LSQR_ATOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**
LSQR_BTOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**
LSQR_CONLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**
LSQR_ITNLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**
LSQRWRITE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0**
* automatic user intervention											
MAXAUI	NA	NA	NA	NA	NA	NA	**87**	**87**	NA	NA	NA
AUISTARTOPT	NA	NA	NA	NA	NA	NA	**1**	**1**	NA	NA	NA
NOAUIPHIRAT	NA	NA	NA	NA	NA	NA	**0.9**	**0.5**	NA	NA	NA
AUIRESTITN	NA	NA	NA	NA	NA	NA	**0**	**0**	NA	NA	NA
AUISENSRAT	NA	NA	NA	NA	NA	NA	**5**	**10**	NA	NA	NA
AUIHOLDMAXCHG	NA	NA	NA	NA	NA	NA	**0**	**0**	NA	NA	NA
AUINUMFREE	NA	NA	NA	NA	NA	NA	**3**	**3**	NA	NA	NA
AUIPHIRATSUF	NA	NA	NA	NA	NA	NA	**0.8**	**0.4**	NA	NA	NA
AUIPHIRATACCEPT	NA	NA	NA	NA	NA	NA	**0.99**	**0.8**	NA	NA	NA
NAUINOACCEPT	NA	NA	NA	NA	NA	NA	**30**	**30**	NA	NA	NA
* parameter groups											
INCTYP	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative
DERINC	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
DERINCLB	0	0	0	0	0	0	0	0	0	0	0
FORCEN	switch	switch	switch	switch	switch	switch	switch	switch	switch	switch	switch
DERINCMUL	2	2	2	2	2	2	2	2	2	2	2
DERMTHD	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic
[SPLITTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SPLITRELDIFF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SPLITACTION]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
* parameter data											
PARTRANS	none	none	none	none	none	None	none	none	none	none	none
PARCHGLIM	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative

Table 2 Continued.

Treatment N°.	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
* control data													
RLAMBDA1	5	5	5	5	5	5	5	5	5	5	5	5	5
RLAMFAC	2	2	2	2	2	2	2	2	2	2	2	2	2
PHIRATSUF	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.1	0.3
PHIREDLAM	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	**0.009**	0.03
NUMLAM	10	10	10	10	10	10	10	10	10	10	10	**20**	10
[JACUPDATE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[LAMFORGIVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[DERFORGIVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
RELPARMAX	3	3	3	3	3	3	3	3	3	3	3	3	3
FACPARMAX	3	3	3	3	3	3	3	3	3	3	3	3	3
FACORIG	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.01
PHIREDSWH	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
[DOAUI]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[DOSENREUSE]	NA	NA	NA	NA	NA	NA	NA	NA	**senreuse**	NA	NA	NA	NA
[BOUNDSCALE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NOPTMAX	50	50	50	50	50	50	50	50	50	50	**100**	**100**	50
PHIREDSTP	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.001**	**0.001**	**0.01**
NPHISTP	3	3	3	3	3	3	3	3	3	3	**6**	**6**	3
NPHINORED	3	3	3	3	3	3	3	3	3	3	**6**	**6**	3
RELPARSTP	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.001**	**0.001**	0.01
NRELPAR	3	3	3	3	3	3	3	3	3	3	**6**	**6**	3
* sensitivity reuse													
SENRELTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	**0.15**	NA	NA	NA	NA
SENMAXREUSE	NA	NA	NA	NA	NA	NA	NA	NA	**-1**	NA	NA	NA	NA
SENALLCALCINT	NA	NA	NA	NA	NA	NA	NA	NA	**3**	NA	NA	NA	NA
SENPREDWEIGHT	NA	NA	NA	NA	NA	NA	NA	NA	**-1**	NA	NA	NA	NA
SENPIEXCLUDE	NA	NA	NA	NA	NA	NA	NA	NA	**yes**	NA	NA	NA	NA
* singular value decomposition													
All variables in this section were as for the baseline													
* lsqr													
LSQRMODE	NA	NA	NA	NA	NA	NA	NA	NA	**1**	NA	NA	NA	NA
LSQR_ATOL	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**	NA	NA	NA	NA
LSQR_BTOL	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**	NA	NA	NA	NA
LSQR_CONLIM	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	NA	NA	NA	NA
LSQR_ITNLIM	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	NA	NA	NA	NA
LSQRWRITE	NA	NA	NA	NA	NA	NA	NA	NA	**0**	NA	NA	NA	NA

* automatic user intervention

All variables in this section were as for the baseline

* parameter groups													
	rel_to_m												
INCTYP	**ax**	relative	relative	relative	relative	relative	Relative	relative	relative	relative	relative	relative	relative
DERINC	0.01	**0.02**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
DERINCLB	0	0	**0.00001**	0	0	0	0	0	0	0	0	0	0
FORCEN	switch	switch	switch	**always_5**	switch	switch	Switch	switch	switch	switch	switch	switch	switch
DERINCMUL	2	2	2	2	**3**	2	2	2	2	2	2	3	2
DERMTHD	parabolic	parabolic	parabolic	**minvar**	parabolic	**best_fit**	Parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic
[SPLITTHRESH	NA	NA	NA	NA	NA	NA	**0.0001**	NA	NA	NA	NA	NA	NA
SPLITRELDIFF	NA	NA	NA	NA	NA	NA	**0.5**	NA	NA	NA	NA	NA	NA
SPLITACTION]	NA	NA	NA	NA	NA	NA	**Smaller**	NA	NA	NA	NA	NA	NA
* parameter data													
PARTRANS	none	none	none	none	none	none	None	**log**	none	none	none	none	none
PARCHGLIM	relative	relative	relative	relative	relative	relative	Relative	**factor**	relative	**factor**	relative	relative	relative

Table 3 Data from PEST control files used to conduct optimisation using Tikhonov regularisation. All optimisation runs were performed using 'regularisation' mode in PEST, with 115 parameters, 56 parameter groups, 69 observation groups and 115 prior equations. Parameters shown in bold indicate variation from the baseline file. PEST control settings in the first column are described in the methods, and parameters in square brackets [] indicate optional use in the PEST control file. Each control file section is identified with an asterisk. NA = Not Applicable.

Treatment N°.	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**
* control data														
All variables in this section were as for the baseline file in Table 2														
* sensitivity reuse														
SENRELTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.15**
SENMAXREUSE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**-1**
SENALLCALCINT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**3**
SENPREDWEIGHT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**-1**
SENPIEXCLUDE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**yes**
* singular value decomposition														
SVDMODE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**2**	NA	NA	NA
MAXSING	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**115**	NA	NA	NA
EIGTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**5E-07**	NA	NA	NA
EIGWRITE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0**	NA	NA	NA
* lsqr														
LSQRMODE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1**	NA	NA
LSQR_ATOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**	NA	NA
LSQR_BTOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**	NA	NA
LSQR_CONLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	NA	NA
LSQR_ITNLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	NA	NA
LSQRWRITE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0**	NA	NA
* automatic user intervention														
All variables in this section were as for the baseline in Table 2														
* parameter groups														
All variables in this section were as for the baseline in Table 2														
* parameter data														
PARTRANS	none	none	none	none	none	none	None	none	none	none	none	none	**log**	**log**
PARCHGLIM	relative	relative	relative	relative	relative	relative	Relative	relative	relative	relative	relative	relative	**factor**	**factor**
* regularisation														
PHIMLIM	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10
PHIMACCEPT	1.05E-10	**1.1E-10**	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10
[FRACPHIM]	0.1	0.1	**0.03**	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
[MEMSAVE]	NA	NA	NA	**memsave**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WFINIT	1	1	1	1	**10**	1	1	1	1	1	1	1	1	1

WFMIN	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10
WFMAX	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10
[LINREG]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[REGCONTINUE]	NA	NA	NA	NA	NA	NA	NA	**continue**	NA	NA	NA	NA	NA	NA
WFFAC	1.3	1.3	1.3	1.3	1.3	**2**	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
WFTOL	0.01	0.01	0.01	0.01	0.01	0.01	**0.001**	0.01	0.01	0.01	0.01	0.01	0.01	0.01
IREGADJ	1	1	1	1	1	1	1	1	**2**	**4**	1	1	1	1
[NOPTREGADJ]	NA	NA	NA	NA	NA	NA	NA	NA	NA	**2**	NA	NA	NA	NA
[REGWEIGHTRAT]	NA	NA	NA	NA	NA	NA	NA	NA	NA	**10**	NA	NA	NA	NA

Table 3 Continued

Treatment No.	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**	**51**	**52[a]**	**53[b]**	**54[c]**
* control data																
RLAMBDA1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
RLAMFAC	2	2	2	2	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	2	**-3**	**-3**	**-3**
PHIRATSUF	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
PHIREDLAM	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
NUMLAM	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
[JACUPDATE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**999**	**999**	NA	NA	NA	NA
[LAMFORGIVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**lamforgive**	**lamforgive**	NA	NA	NA	NA
[DERFORGIVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**derforgive**	**derforgive**	NA	NA	NA	NA
RELPARMAX	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
FACPARMAX	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
FACORIG	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
PHIREDSWH	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
[DOAUI]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[DOSENREUSE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[BOUNDSCALE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NOPTMAX	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
PHIREDSTP	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
NPHISTP	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
NPHINORED	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
RELPARSTP	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
NRELPAR	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
* Sensitivity reuse																
SENRELTHRESH	**0.15**	NA	**0.15**	NA	NA	**0.15**	NA	**0.15**	NA	**0.15**	**0.15**	NA	NA	NA	NA	NA
SENMAXREUSE	**-1**	NA	**-1**	NA	NA	**-1**	NA	**-1**	NA	**-1**	**-1**	NA	NA	NA	NA	NA
SENALLCALCINT	**3**	NA	**3**	NA	NA	**3**	NA	**3**	NA	**3**	**3**	NA	NA	NA	NA	NA

SENPREDWEIGHT	**-1**	NA	**-1**	NA	NA	**-1**	NA	**-1**	NA	**-1**	**-1**	NA	NA	NA	NA	NA
SENPIEXCLUDE	**yes**	NA	**yes**	NA	NA	**yes**	NA	**yes**	NA	**yes**	**yes**	NA	NA	NA	NA	NA
* singular value decomposition																
SVDMODE	**2**	**2**	NA	NA	NA	NA	**2**	**2**	NA	NA	**2**	**2**	NA	NA	NA	NA
MAXSING	**115**	**115**	NA	NA	NA	NA	**115**	**115**	NA	NA	**115**	**115**	NA	NA	NA	NA
EIGTHRESH	**5E-07**	**5E-07**	NA	NA	NA	NA	**5E-07**	**5E-07**	NA	NA	**5E-07**	**5E-07**	NA	NA	NA	NA
EIGWRITE	**0**	**0**	NA	NA	NA	NA	**0**	**0**	NA	NA	**0**	**0**	NA	NA	NA	NA
* lsqr																
LSQRMODE	NA	NA	**1**	**1**	**1**	**1**	NA	NA	**1**	**1**	NA	NA	NA	**1**	**1**	**1**
LSQR_ATOL	NA	NA	**0.0001**	**0.0001**	**0.0001**	**0.0001**	NA	NA	**0.0001**	**0.0001**	NA	NA	NA	**0.0001**	**0.0001**	**0.0001**
LSQR_BTOL	NA	NA	**0.0001**	**0.0001**	**0.0001**	**0.0001**	NA	NA	**0.0001**	**0.0001**	NA	NA	NA	**0.0001**	**0.0001**	**0.0001**
LSQR_CONLIM	NA	NA	**1000**	**1000**	**1000**	**1000**	NA	NA	**1000**	**1000**	NA	NA	NA	**1000**	**1000**	**1000**
LSQR_ITNLIM	NA	NA	**1000**	**1000**	**1000**	**1000**	NA	NA	**1000**	**1000**	NA	NA	NA	**1000**	**1000**	**1000**
LSQRWRITE	NA	NA	**0**	**0**	**0**	**0**	NA	NA	**0**	**0**	NA	NA	NA	**0**	**0**	**0**
* automatic user intervention																
All variables in this section were as for the baseline in Table 2																
* parameter groups																
INCTYP	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative	relative
DERINC	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
DERINCLB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FORCEN	switch	switch	switch	switch	Switch	switch	switch	switch	switch	switch	switch	switch	switch	switch	switch	switch
DERINCMUL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
DERMTHD	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic	parabolic
[SPLITTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.0001**	**0.0001**	NA	NA	NA	NA
SPLITRELDIFF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.5**	**0.5**	NA	NA	NA	NA
SPLITACTION]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**smaller**	**smaller**	NA	NA	NA	NA
* parameter data																
PARTRANS	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	none	none	**log**	**log**	none	**log**	**log**	**log**
PARCHGLIM	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	relative	relative	**factor**	**factor**	relative	**factor**	**factor**	**factor**
* regularisation																
PHIMLIM	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10
PHIMACCEPT	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10	1.05E-10
[FRACPHIM]	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
[MEMSAVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WFINIT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
WFMIN	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10	1E-10
WFMAX	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10	1E+10
[LINREG]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**linreg**	NA	NA	NA
[REGCONTINUE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WFFAC	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3

WFTOL	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
IREGADJ	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
[NOPTREGADJ	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[REGWEIGHTRAT]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

[a] Increased weighting on all N_2O measurements from 4.203 (as applied in all other treatments) to 20.

[b] Adjusted weighting on all measured data such that the contribution to ϕ from each observation group was 10.

[c] Repeated optimisation using optimised APSIM parameters from treatment 43.

Table 4 PEST control file settings used to conduct global optimisation with CMAES_P. All optimisation runs were performed using 'estimation' mode in PEST, with 84 parameters, 39 parameter groups, 13 observation groups and zero prior equations. Parameters shown in bold indicate variation from the baseline file in Table 2. CMAES_P control settings in the first column are described in the methods, with values in round brackets () in the CMAES_P section of the table indicating the default value for CMAES_P. Parameters in square brackets [] indicate optional use in the PEST control file. Each control file section is identified with an asterisk. NA = Not Applicable, N = no, Y = yes, S = superlinear, L = linear, E = equal.

Treatment no	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**
* control data																	
All parameters in this section except for DOSENREUSE were as for the baseline in Table 2																	
[DOSENREUSE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**senreuse**	NA
* Sensitivity reuse																	
SENRELTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.15**	NA
SENMAXREUSE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**-1**	NA
SENALLCALCINT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**3**	NA
SENPREDWEIGHT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**-1**	NA
SENPIEXCLUDE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**yes**	NA
* singular value decomposition																	
SVDMODE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**2**	NA	NA	**2**	NA
MAXSING	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**115**	NA	NA	**115**	NA
EIGTHRESH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**5.00E-07**	NA	NA	**5.00E-07**	NA
EIGWRITE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0**	NA	NA	**0**	NA
* lsqr																	
LSQRMODE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1**	**1**	NA	NA
LSQR_ATOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E-04**	**1.00E-04**	NA	NA
LSQR_BTOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E-04**	**1.00E-04**	NA	NA
LSQR_CONLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	**1000**	NA	NA
LSQR_ITNLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1000**	**1000**	NA	NA
LSQRWRITE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0**	**0**	NA	NA
* automatic user intervention																	
All variables in this section were as for the baseline in Table 2																	
* parameter groups																	
All variables in this section were as for the baseline in Table 2																	
* parameter data																	
PARNME	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PARTRANS	none	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	**log**	none

PARCHGLIM	relative	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	**factor**	relative
* regularisation																	
PHIMLIM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E-10**	**1.00E-10**	**1.00E-10**	**1.00E-10**	**1.00E-10**	NA
PHIMACCEPT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.05E-10**	**1.05E-10**	**1.05E-10**	**1.05E-10**	**1.05E-10**	NA
[FRACPHIM]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	NA
[MEMSAVE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WFINIT	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1**	**1**	**1**	**1**	**1**	NA
WFMIN	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E-10**	**1.00E-10**	**1.00E-10**	**1.00E-10**	**1.00E-10**	NA
WFMAX	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E+10**	**1.00E+10**	**1.00E+10**	**1.00E+10**	**1.00E+10**	NA
[LINREG]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[REGCONTINUE]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
WFFAC	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.3**	**1.3**	**1.3**	**1.3**	**1.3**	NA
WFTOL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1.00E-02**	**1.00E-02**	**1.00E-02**	**1.00E-02**	**1.00E-02**	NA
IREGADJ	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**1**	**1**	**1**	**1**	**1**	NA
[NOPTREGADJ]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[REGWEIGHTRAT]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[REGSINGTHRESH]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Variables within the CMAES_P algorithm via the command prompt																	
Populations size (ψ = 4 + 3*ln(n))	18	18	18	18	18	18	**25**	18	18	18	**25**	18	18	18	**25**	**25**	**25**
Number of parents (ω = ψ/2)	9	9	9	9	9	9	9	**4**	**15**	9	9	9	9	9	9	9	9
Recombination weights (superliner, linear or equal)	s	s	s	s	s	**E**	s	s	s	s	s	s	s	s	s	s	s
SVD-hybridization (N)	N	N	**Y**	**Y**	**Y**	N	N	N	N	N	N	N	N	N	N	N	N
Singular value trial threshholds (3)	NA	NA	**3**	**3**	**10**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Hybridization (soft)	NA	NA	**hard**	**soft**	**soft**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Forgive model run failure (Y)	Y	Y	Y	Y	Y	Y	Y	Y	Y	**N**	**N**	Y	Y	Y	**N**	**N**	Y

Results

Using PEST to parameterise APSIM resulted in significant improvement in model fit (as shown by comparing the ϕ value from Treatment 0 to that from other treatments in Table 5), but this was mainly because part of the parameterisation process required finding optimal plant parameters for a maize hybrid (PR34N43) that was not included in the default APSIM release. Parameterisation with the default GML algorithm (and PEST control file settings) reduced ϕ from 244 (no calibration, treatment 0) to ~82 or less, with the corresponding correlation coefficients increasing from 0.37 to 0.81 or greater (Table 5). Both the value of ϕ and that of r indicate that PEST improved the predictability of APSIM for the scenario modelled. For the default GML algorithm, ϕ ranged from 82.3 (treatments 12-15) to 28.8 (treatments 11 and 20). Parameterisation routines employing SVD or LSQR resulted in lower ϕ values than all of the other parameterisation options we examined here. However, algorithms that used LSQR were completed significantly earlier than those using SVD (Table 5), indicating that the former would be a better choice for model parameterisation. Correlation coefficients generally improved as the value of ϕ decreased, indicating better fit.

Table 5 shows that LSQR was one of the faster performing algorithms, and that reuse of composite parameter sensitivity (SENREUSE) generally reduced optimisation times even further, without detriment to the quality of the fit. As a consequence, many of the Tikhonov regularisation optimisation settings we used in Table 3 were in conjunction with the SENREUSE setting.

Twelve treatments that used regularisation resulted in lower ϕ values than those without regularisation. Treatment 52 examined the effect of reusing the optimised parameters from one of the better performing treatments (treatment 43, which included logarithmic transformation, LSQR and setting RLAMFAC to -3), and increasing the weight applied to the N_2O dataset, since this dataset generally carried the lowest ϕ (Table 6). Treatment 53 was also similar to treatment 43, except weights of each observation group (grain yield, N_2O etc.) in treatment 53 were adjusted such that based on initial parameter values, the contribution made to ϕ by each observation group was approximately 10. The higher ϕ values of treatments 52 and 53 indicate that either reusing optimised parameters and increasing weight assigned to observations groups that dominate the value of ϕ or equalising the weighting of each observation groups prior to optimisation using PEST's PWTADJ1 utility program reduce the quality of optimisation, as shown by comparison of these treatments to their equivalent treatments with no reweighting of observation groups (treatments 54 and 43, respectively).

The PEST control file setting FRACPHIM (treatment 27) had little effect on ϕ. Although log transformation of variables prior to inversion was useful in lowering ϕ when used in concert with SVD and LSQR, log transformation of parameters alone did not improve the quality of fit when used with regularisation (treatment 37), or with composite parameter sensitivity reuse (SENREUSE; treatment 38), or in treatments that combined SENREUSE, LSQR, and/or an RLAMFAC value of -3 (treatments 41, 44 in Table 6).

As with the default optimisation algorithms, there was no apparent relationship between quality of fit and optimisation time, with optimisation times varying from 4,684 to 44,238 (Table 6). Optimisation settings in combination with regularisation that afforded the best fits included the use of LSQR or SVD, logarithmic transformation of variables, and (if altered from the default value of 2), setting RLAMFAC to -3. Although SENREUSE generally resulted in lower run times, evidence from comparisons of treatments 38 to 37, 49 to 50, 46 to 45 and 42 to 41 indicated that SENREUSE generally caused premature termination of the optimisation algorithm, resulting in some loss of quality of model parameterisation when used with regularisation.

Table 5 CPU time, value of the objective function (ϕ) and Pearson's correlation coefficient (r) resulting from optimisation of APSIM parameters using PEST. Optimisation iterations were performed using different settings in the PEST control file (see methods and Table 1). Shaded rows represent no optimisation (treatment 0) or PEST default optimisation settings. Rows are arranged with ϕ in descending order.

Treatment Nº.	PEST control file settings examined	CPU time (s)	ϕ	r
0	NONE	NA	244	0.37
12	REL TO MAX	4172	82.3	0.81
14	DERINCLB	2557	82.3	0.81
15	ALWAYS 5	10696	82.2	0.81
13	DERINC	2474	82.1	0.81
23	PHIREDLAM	12386	81.6	0.81
22	NOPTMAX	11557	81.6	0.81
4	NPHISTP	7922	81.6	0.81
24	FACORIG	2972	81.6	0.82
17	BEST FIT	2716	81.6	0.81
18	SPLITTHRESH	2706	81.6	0.81
16	DERINCMUL	2692	81.6	0.81
1	DEFAULT	2595	81.6	0.81
6	SENREUSE	1478	81.6	0.81
2	RLAMBDA	4614	81.4	0.81
5	LAMFORGIVE	4316	79.6	0.82
3	RELPARMAX	5918	79.4	0.82
7	AUID	4267	79.0	0.82
21	FACTOR	4776	78.5	0.82
19	LOG	4449	77.5	0.82
9	SVD, BOUNDSCALE	6796	73.4	0.81
8	AUI, SENREUSE	3996	32.7	0.92
10	SVD	11611	28.9	0.93
20	LSQR, SENREUSE	2555	28.8	0.93
11	LSQR	3380	28.8	0.93

Table 6 CPU time, objective function (ϕ) and Pearson's correlation coefficient (r) from optimisation of APSIM parameters using PEST with Tikhonov regularisation (REG). The shaded row represents optimisation results conducted using PEST default settings with regularisation. Other treatments were conducted using regularisation and different settings in the PEST control file (see methods and Table 1). Rows are arranged with ϕ in descending order.

Treatment Nº.	PEST control file settings examined	CPU time (s)	ϕ	r
52	REG, LOG, LSQR, RLAMFAC, PARREP, N2OWT	7831	380	0.71
53	REG, LOG, LSQR, RLAMFAC, PWTADJ1	10998	100	0.74
27	REG, FRACPHIM	13851	82.3	0.81
37	REG, LOG	15935	36.5	0.91
51	REG, LINREG	27787	35.8	0.91
31	REG, WFTOL	21481	35.8	0.91
38	REG, LOG, SENREUSE	11761	35.4	0.91
41	REG, LOG, LSQR, SENREUSE	4850	35.1	0.91
44	REG, LOG, LSQR, RLAMFAC, SENREUSE	4684	35.1	0.91
33	REG, IREGADJ	34942	34.1	0.92
32	REG, REGCONTINUE	34929	34.1	0.92
25	REG	29493	34.1	0.92
26	REG, PHIMACCEPT	28367	34.1	0.92
30	REG, WFFAC	17082	33.9	0.92
29	REG, WFINIT	39968	31.3	0.93
49	REG, LOG, SVD, RLAMFAC, SPLITTHRESH, LAMFORGIVE, SENREUSE	12889	30.3	0.93
35	REG, SVD	20222	28.9	0.93
48	REG, LSQR, RLAMFAC, SENREUSE	8238	28.9	0.93
36	REG, LSQR	27136	28.6	0.93
47	REG, LSQR, RLAMFAC	22393	28.5	0.93
28	REG, MEMSAVE	24124	28.4	0.93
34	REG, IREGADJ4	30843	28.1	0.93
39	REG, LOG, SVD, SENREUSE	11338	28.1	0.93
50	REG, LOG, SVD, RLAMFAC, SPLITTHRESH, LAMFORGIVE	14632	27.5	0.93
40	REG, LOG, SVD	15975	27.4	0.93
42	REG, LOG, LSQR	25862	26.8	0.93
46	REG, LOG, SVD, RLAMFAC, SENREUSE	21985	26.5	0.93
45	REG, LOG, SVD, RLAMFAC	18965	26.4	0.93
54	REG, LOG, LSQR, RLAMFAC, PARREP	13262	26.2	0.93
43	REG, LOG, LSQR, RLAMFAC	44238	26.0	0.93

Table 7 CPU time, objective function (ϕ) and Pearson's correlation coefficient (r) resulting from optimisation of APSIM parameters using the PEST algorithm CMAES_P (covariance matrix adaptation evolution strategies). The shaded row represents optimisation results conducted using default CMAES_P settings. Other treatments were conducted using CMAES_P with different settings in the PEST control file (see methods and Table 1). Rows are arranged with ϕ in descending order. PE = prior equations (see methods for descriptions of CMAES_P simulations).

Treatment N°.	CMAES_P settings examined	CPU time (s)	ϕ	r
67	LOG, PE, SVD	22465	47.7	0.88
68	LOG, PE, LSQR	16693	47.7	0.88
63	LOG, OMEGA15	19141	39.2	0.90
62	LOG, OMEGA4	18341	36.9	0.91
60	LOG, EQUAL WT	61409	36.6	0.91
57	LOG, SVDHYBR	33878	31.2	0.92
55	DEFAULT CMAES	47749	30.5	0.92
66	LOG, PE	76195	28.8	0.93
71	PSI	62947	25.8	0.94
70	LOG, PSI, NOFORGIVE, PE, SVD, SENREUSE	92370	24.4	0.94
59	LOG, SVDHYBR, SOFT10	32941	23.0	0.95
58	LOG, SVDHYBR, SOFT	31698	23.0	0.95
56	LOG	58469	18.5	0.97
64	LOG, NOFORGIVE	55214	18.5	0.97
61	LOG, PSI	56644	15.9	0.98
65	LOG, PSI, NOFORGIVE	47484	15.9	0.98
69	LOG, PSI, NOFORGIVE, LSQR, RLAMFAC	47243	15.9	0.98

Parameterisation using CMAES_P resulted in the lowest ϕ values of all optimisation algorithms examined (Table 7). These values were obtained by log transforming variables and increasing the initial population size (ψ) to 25 (from a default population size of 18 in treatment 55). When failure of CMAES_P to read any part of a model run was allowed to precipitate cessation of the algorithm (i.e. setting NOFORGIVE in Table 7 to true), total run times were generally reduced (cf. treatment 64 to 56 and 65 to 61). This occurred because in the default case, parameter vectors giving rise to model run failure did not terminate the CMAES_P algorithm. Instead, the default case allowed CMAES_P to continue after internally assigning a very high objective function to this parameter set, thus providing a disincentive to the optimisation process from generating similar parameter sets in future runs (Doherty, 2016a).

Treatments 61, 65 and 69 resulted in the lowest ϕ of all CMAES_P algorithms examined (all used log transformation and increased the population size to 25); treatments 65 and 69 terminated optimisation earlier due to the NOFORGIVE setting detailed above. Although ϕ values for many CMAES_P treatments were lower than those obtained using regularisation (Table 6), CMAES_P optimisation run times were considerably longer.

The shortest optimisation time of all CMAES_P treatments was obtained using log transformation of parameters, prior equations, and LSQR (treatment 68), although the value of ϕ associated with this treatment was relatively high. The longest optimisation times were obtained by increasing the value of population size, log transforming variables, or both. Treatment 70 took the longest period for parameter convergence, requiring over 25 hrs of CPU time. This treatment included log transformation of variables, increased population size, use of prior equations, SVD, composite parameter sensitivity reuse, and the NOFORGIVE setting detailed above. Treatments 66 (log transformation of variables and use of prior equations), 71 (increased population size) and 60 (log transformation and equally weighting all of the parent population vectors) also required extensive CPU times to satisfy the parameter convergence criteria applied here.

The relationship between ϕ and the total number of model calls shown in Fig. 1 depicts the trade-off between lower cumulative residual error (between modelled and measured values) with CPU time; generally lower ϕ values were obtained after more model calls. Optimisation runs with the default solver or Tikhonov regularisation required the fewest model calls for parameter convergence, but the majority of regularisation runs resulted in lower ϕ values than those using the default solver method.

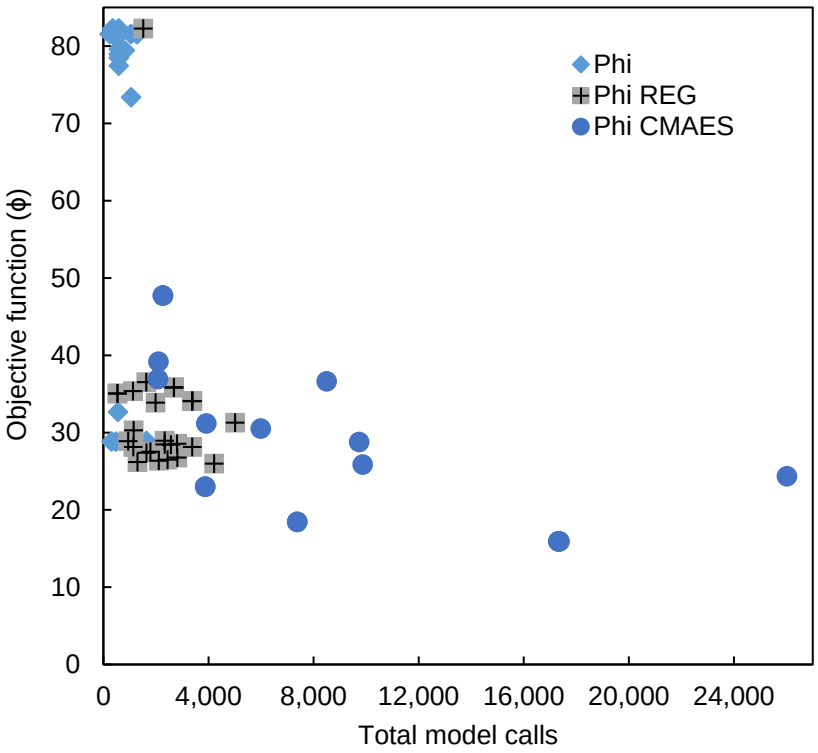


Fig. 1. Relationship between the total number of model calls with the objective function (ϕ) of each optimisation run in PEST (treatments 52 and 53 not shown since weighting applied to objective function components in these treatments was not comparable to other treatments).

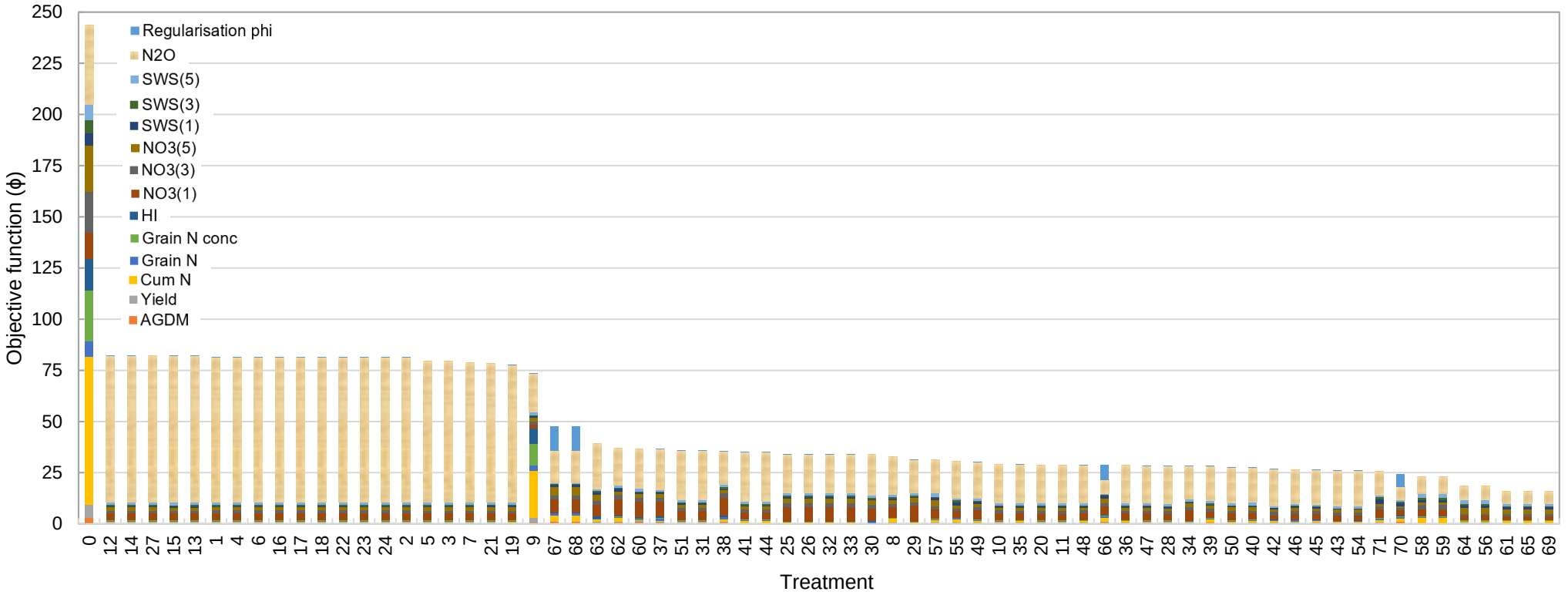
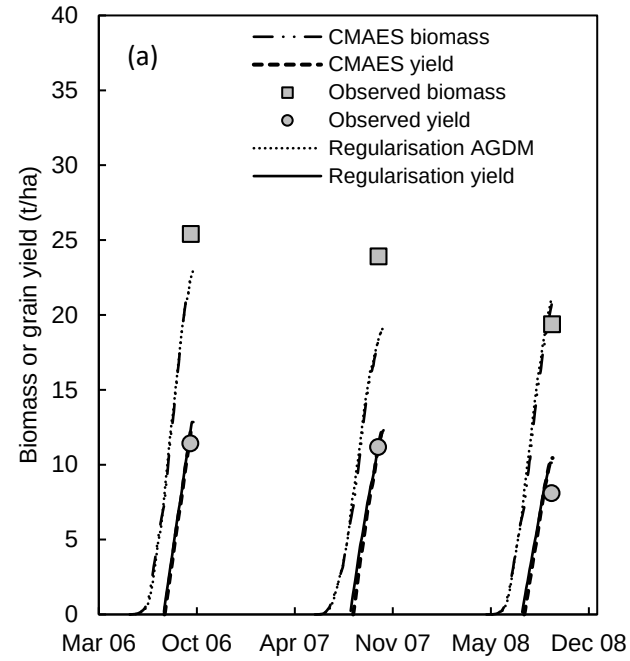
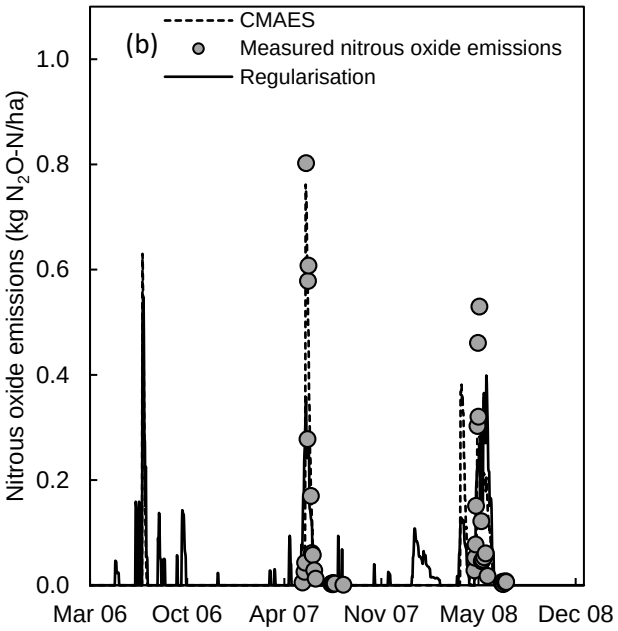


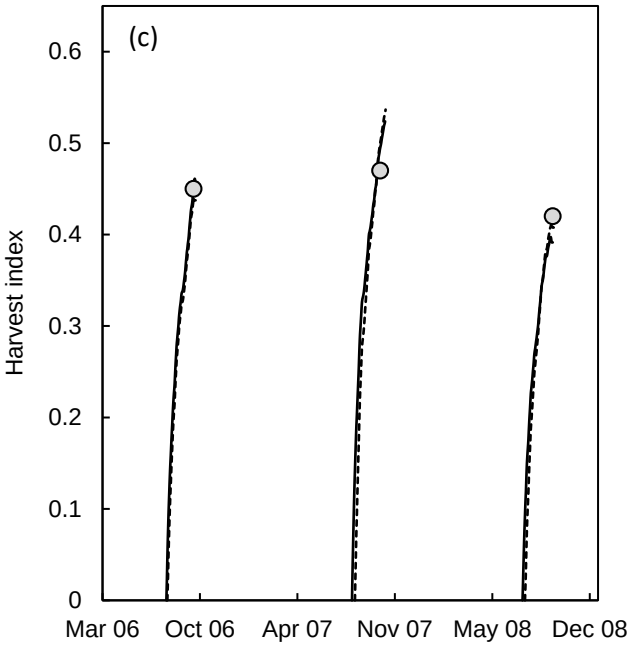
Fig. 2. Contribution of each APSIM variable to the total objective function value (ϕ) for each treatment in descending order from left to right. Legend abbreviations: AGDM = above-ground dry matter, Grain N = grain nitrogen content in kg/ha, HI = harvest index, Grain N conc = N concentration of grain in %, N_2O = nitrous oxide emissions in kg N_2O-N/ha, Cum N = cumulative N uptake in above-ground biomass in kg/ha, NO3(1, 3, 5) = nitrate content in kg/ha in soil layers 1, 3 or 5, respectively, SWS(1, 3, 5) = soil water content in layers 1, 3 and 5 respectively, yield = grain yield, and regularisation phi = contribution of prior equations to ϕ.

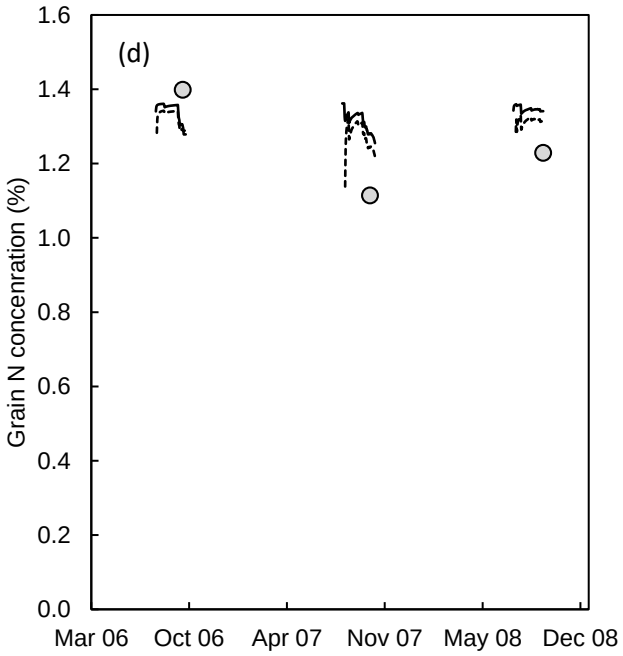
Variability in ϕ across CMAES_P runs was generally higher than that for other optimisation methods (Fig. 1). The clustering of default and regularisation optimisation methods around a ϕ value of 26 indicates a local minimum in the objective function surface at this point, since some of the CMAES methods were able to attain ϕ values as low as 15.9.

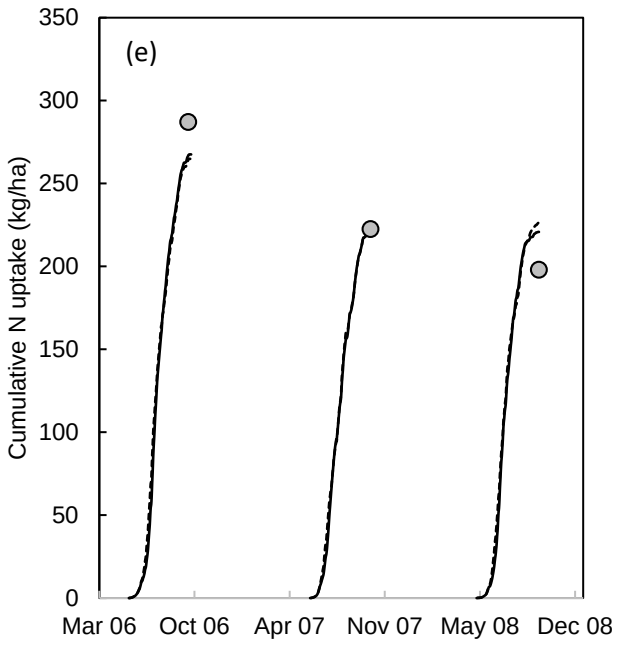
The contribution to ϕ from each variable for each treatment is shown in Fig. 2. The uncalibrated treatment (0) had the highest ϕ value, mainly due to large error in the prediction of N_2O, soil nitrate, grain N concentration and cumulative N uptake. Treatments conducted with the default GML algorithm had the highest ϕ values (e.g. 12, 14, 27 and 15). Treatments 61, 65 and 69 had the lowest ϕ values (as discussed above); each of these treatments were conducted with CMAES_P. CMAES_P treatments were most effective at lowering ϕ due to their ability to reduce model error associated with nitrous oxide emissions.

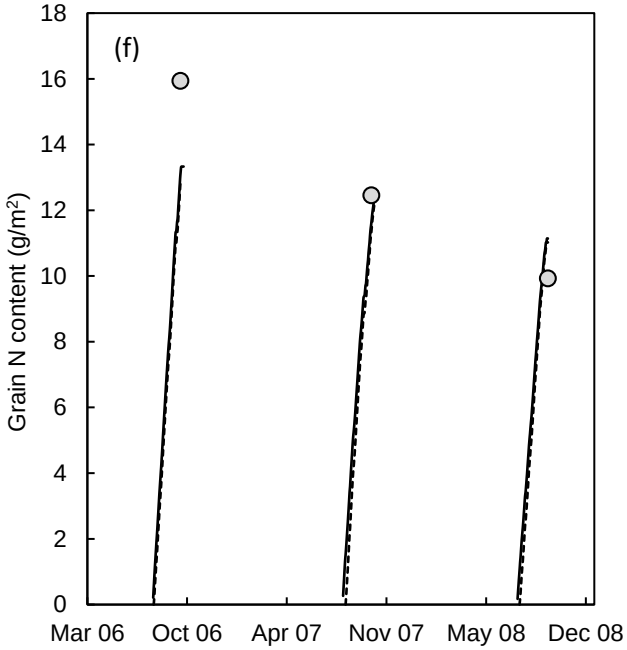

(a)
CMAES biomass
CMAES yield
Observed biomass
Observed yield
Regularisation AGDM
Regularisation yield
Biomass or grain yield (t/ha)
0
5
10
15
20
25
30
35
40
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08

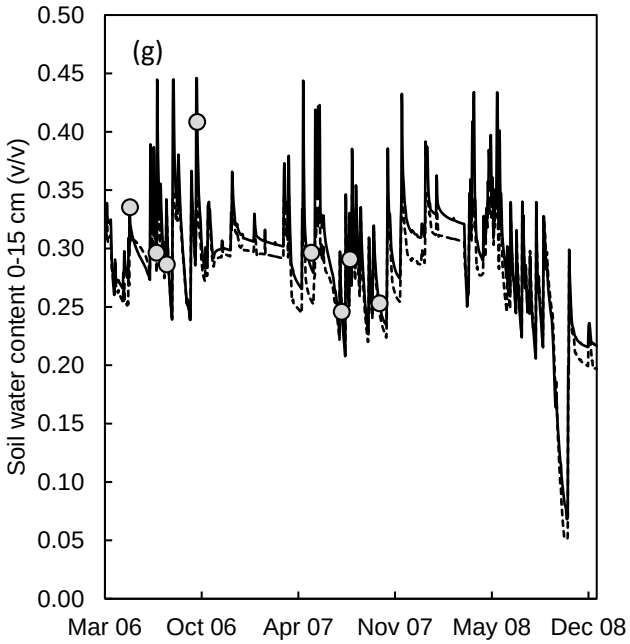

(b)
CMAES
Measured nitrous oxide emissions
Regularisation
Nitrous oxide emissions (kg N2O-N/ha)
0.0
0.2
0.4
0.6
0.8
1.0
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08

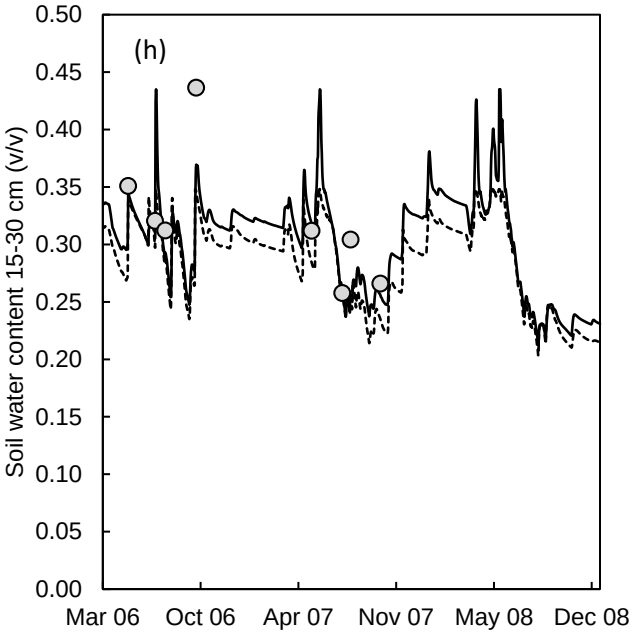

(c)
Harvest index
0
0.1
0.2
0.3
0.4
0.5
0.6
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08


(d)
Grain N concenration (%)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08


(e)
Cumulative N uptake (kg/ha)
350
300
250
200
150
100
50
0
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08


(f)
Grain N content (g/m²)
18
16
14
12
10
8
6
4
2
0
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08


(g)
Soil water content 0-15 cm (v/v)
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08


(h)
Soil water content 15-30 cm (v/v)
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Mar 06
Oct 06
Apr 07
Nov 07
May 08
Dec 08

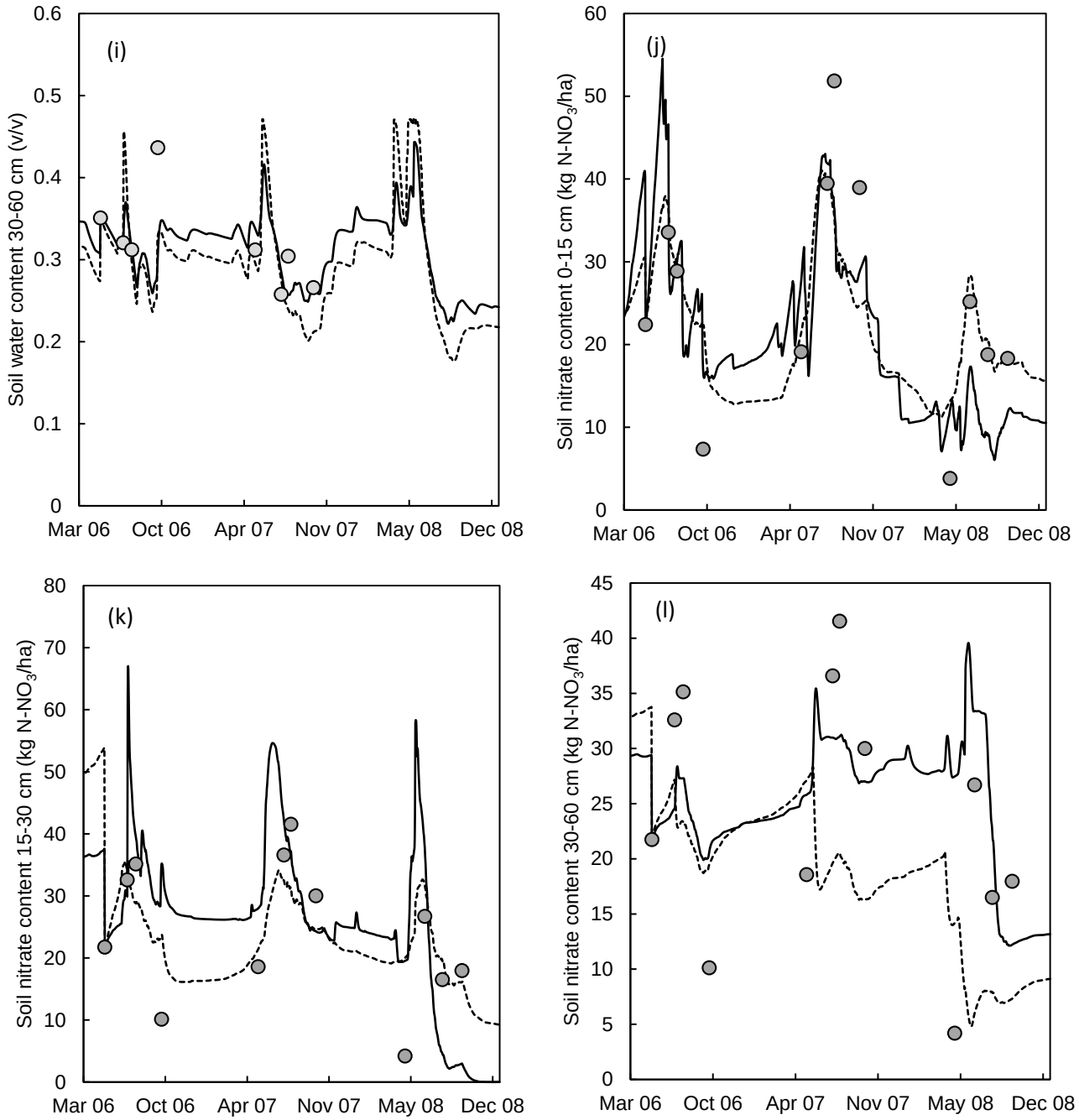


Fig. 3. Measured (points) and simulated (lines) **(a)** above-ground biomass and grain yield, **(b)** soil nitrous oxide emissions, **(c)** harvest index, **(d)** grain nitrogen concentration, **(e)** cumulative N uptake, **(f)** grain N content, **(g)** volumetric soil water content in the first layer (0-15 cm), **(h)** volumetric soil water content in the second layer (15-30 cm), **(i)** volumetric soil water content in the third layer (30-60 cm), **(j)**, soil nitrate content in the first layer (0-15 cm), **(k)** soil nitrate content in the second layer (15-30 cm), **(l)** soil nitrate content in the third layer (30-60 cm). The solid line represents treatment 43, whilst the dashed lines represent treatment 65 (see Tables 3 and 4 for features of these treatments).

In general, the greatest contribution to ϕ was from N_2O, except for treatment 9, which employed BOUNDSCALE and SVD (as discussed above). The contribution to ϕ from the regularisation objective function (ϕ_r in Eqn. 6) in applicable treatments was generally very small; the largest contribution to ϕ_r was in treatments 70, 66, 67 and 68. Treatments that employed Tikhonov regularisation also had greater contributions from NO_3 and soil water content to ϕ.

Two of the better performing PEST optimisation settings are shown for comparison in Fig. 3. These included treatment 43 (regularisation, logarithmic transformation of parameters prior to inversion, LSQR and setting RLAMFAC to -3), and treatment 65 (CMAES with logarithmic transformation of parameters, an initial population size of 25, and terminating the optimisation whenever any parameter upgrade vector resulted in model run failure). Differences in the quality of the fit to either grain yield of biomass were minimal over the three years (Fig. 2a). Treatment 65 provided a better fit to N_2O data in 2007, where the peak N_2O emissions in that year were greater, but in 2008 the timing of the simulation parameterised by treatment 43 was more realistic than that provided by treatment 65 (Fig. 3b). Differences between the quality of parameterisation of harvest index, grain N concentration, cumulative N uptake and grain N content were minimal (Figs 3c-f). In most cases, the parameterisation by treatment 65 resulted in lower soil water content in each of the three layers, and slightly larger differences between measured values and simulations. Treatment 65 resulted in a parameter set that caused more damping of the temporal fluctuations in NO_3 content, and slightly lower ARMSE overall for soil NO_3 content (Fig. 2). This resulted in a better fit to observed NO_3 in 2008 in the first and second soil layers (0-15 cm, 15-30 cm), but not necessarily in other years or in the third layer.

Discussion

Use of auto-parameterisation as a tool for model intercomparison and identifiability analysis

The purpose of this study was to examine a range of PEST optimisation algorithms on the biophysical systems model, APSIM (Keating et al., 2003). Although PEST has been used previously on APSIM (Akponikpè et al., 2010; Chen et al., 2016), the extent to which the alternative optimisation algorithms within PEST could minimise the weighted sum of squared residuals between the measured and modelled data was unknown. Indeed, past studies using PEST with APSIM have predominantly had agronomic foci, whereas the lens of this study was on optimisation algorithms. This study has demonstrated that several combinations of optimisation algorithms within PEST can be reliably used to perform multi-objective function optimisation of APSIM.

Aside from specification of preferred parameterisation settings in PEST, the observation that auto-parameterisation can be used in this way to calibrate crop models is a very useful finding *per se*. There have been several past agricultural studies that have compared the outputs across models (Rosenzweig et al., 2014; Rosenzweig et al., 2013). During such model inter-comparison studies, different users were required to calibrate their model to intensive time-series datasets that were typically measured in the field (Rosenzweig et al., 2013). However, the extent to which anthropogenic effects influence differences between results of such studies is unknown. In particular, non-uniform prioritisation of variables used in the calibration process may lead to differences in model predictions purely as a result of the philosophical basis that a user places in different variables as part of the calibration process. If PEST (or other optimisation algorithm or software) was consistently used in such inter-comparison studies, one could argue that much of the user subjectivity might be removed by selecting the same optimisation algorithm and termination criteria for model parameterisation. The variability associated with the model user *per se* would then be the selection of lower and upper bounds placed on parameters, as well as initial parameter values. Where more than one team used the same model in an inter-comparison study, common initial values and bounds could be placed on model parameters. The same approach could be adopted in cases where a given parameter was common in multiple models (e.g. leaf N concentrations, radiation-use efficiency, specific leaf area, etc.). The important point here is that if

auto-parameterisation were able to absolve much of the uncertainty associated with how users parameterise models, the remaining differences between modelled results should better reflect model structural differences and thus strengths and weaknesses in model subroutines.

Automated calibration approaches such as this study can also be used with identifiability analyses (Doherty and Hunt, 2009). Identifiability analyses evaluate the degree to which parameters can be estimated uniquely by relating the contributions made by the adjustable parameters to any of the eigenvectors spanning the calibration solution space (Necpálová et al., 2015). Unlike sensitivity analysis, identifiability analysis accounts for parameter correlations that can make it impossible to uniquely estimate even highly sensitive model parameters (Doherty and Hunt, 2010). In conjunction with automated calibration approaches, identifiability analyses have been used to show that only a small number of parameters used in most environmental models are uniquely estimable with most datasets (Beck and Halfon, 1991). Inability to uniquely identify model parameters can be the result of their high correlation with other parameters, or lack of sensitivity of the model outputs to these parameters (Necpálová et al., 2015). This sort of problem is very difficult to recognise without specialised tools and can lead to misidentification of parameter values and inaccurate model projections for conditions outside the range of the calibration dataset. The application of inverse modelling provides insights about parameter dependencies, parameters that exert the greatest influence on the simulated values, whether field observations contain sufficient information to estimate the model parameters, and uncertainty associated with the predictions based on the estimated parameter values (Necpálová et al., 2015).

Preferred optimisation settings in PEST

Implementing Tikhonov regularisation was occasionally an improvement upon the GML algorithm without Tikhonov regularisation when SVD or LSQR was also employed (cf. Tables 5 and 6). Of the 30 regularisation treatments examined, only 12 had lower ϕ than that obtained from GML treatments 11, 20 and 10, which were conducted with SVD or LSQR (or LSQR with parameter sensitivity reuse; Table 5). The 12 regularisation treatments with lowest ϕ most often included LSQR or SVD and of these, the eight treatments with the lowest ϕ had log transformed parameters prior to inversion. In general, however, Pearson's correlation coefficient values for treatments 11, 20, 10 and all of the aforementioned regularisation treatments were greater than 0.9, suggesting any of these treatments - regardless of many other PEST control file settings - provided a high degree of parameterisation. Moreover, the majority of the treatments without regularisation converged much earlier than those with regularisation, indicating that if parameterisation run time is an issue, users may simply opt to apply the default algorithm with SVD or LSQR.

Somewhat counter to our expectations, regularisation treatments with SVD were often terminated faster than corresponding LSQR treatments (e.g. treatment 43 required 44,238 seconds and SVD treatment 45 required 18,965 seconds; treatment 42 required 25,862 seconds whereas the corresponding SVD treatment (40) converged in 15,975 seconds; Table 6). Doherty and Hunt (2010) suggest that the computational costs when employing SVD arise from calculation of the Jacobian matrix and linearisation of the solution in the search direction, such that computational time can become inordinately large when more than 2,500 parameters are optimised (Lin et al., 2016). In contrast to SVD, however, LSQR approximates the solution using a least squares subspace, where the algorithm projects the original problem down to a subspace and solves the projected problem, instead of finding the solution in the original parameter subspace. This projection usually results in much smaller dimensionality of the problem and thus reduced computation costs (Lin et al., 2016).

Although we found evidence of this for the default GML algorithm without regularisation (cf. treatments 10 and 11 in Table 5), the opposite was true when regularisation was employed (Table 6). This may have been because we estimated much fewer than 2,500 parameters or because implementation of SVD or LSQR with Tikhonov regularisation increased the computational burden due to incorporation of prior information equations.

One of the better performing regularisation algorithms included log transformation of parameters, LSQR, and setting the factor used to adjust λ between successive iterations (RLAMFAC) to -3 (treatment 43). Reusing optimised parameters from this treatment in a subsequent round of optimisation did not improve the quality of the fit (treatment 54), indicating that the optimisation algorithm had converged on a local or global minimum in treatment 43. For all regularisation treatments, treatment 43 resulted in the lowest ϕ for optimisation using Tikhonov regularisation, but required 4,216 model calls and thus was one of the longest running regularisation methods examined. CMAES_P treatments 61, 65 and 69 resulted in even better fits to the data (these treatments were conducted with log transformed variables prior to optimisation and increased initial population size (ψ)), but required even more model calls (17,317-17,357) than the regularisation methods resulting in the lowest ϕ (treatments 54 and 43). These results suggest that (1) log transformation of parameters prior to inversion, (2) implementing LSQR and (3) setting RLAMFAC to -3 appear to be better optimisation settings with APSIM if run time is an issue. Tikhonov regularisation is clearly a faster optimisation method than CMAES_P, likely because CMAES_P runs require multiple (ψ) vector upgrades during every iteration, whereas the GML method with Tikhonov regularisation does not. Algorithms combining log transformation of parameters, SVD, regularisation (and potentially an RLAMFAC of -3) also performed well with respect to terminal values of ϕ (e.g. treatments 40, 46 and 45 in Table 6). As the four regularisation treatments with lowest ϕ were realised when RLAMFAC was set to -3 (from the default value of 2), our advice to future research on this theme is a thorough examination of how RLAMFAC affects parameter convergence and optimisation time.

Here we implemented prior equations implicit to Tikhonov regularisation using the ADDREG1 utility program provided with the PEST suite. This program formulated prior information equations as linear equations involving individual parameters. As such, prior information equations did not involve more than one parameter, although multiple parameters can be included in such equations so long as the overall equation is linear. Previous studies that have used Tikhonov regularisation to solve optimisation problems have shown that whilst the performance of the algorithm is not strongly dependent on the prior knowledge equations, a preliminary estimate of parameter values (in the form of prior equations) enables more accurate estimation of parameters (Rouchier et al., 2015). Experimentation with different initial parameter values and thus prior information is something that should be explored in future studies.

To assess whether prior information was useful in global optimisation, we added the same prior information from the regularisation control files to some of the CMAES_P files. However, in contrast to Tikhonov regularisation, use of prior information with CMAES_P resulted in little improvement (treatment 66) or degraded the quality of the optimisation (treatments 67 and 68), the latter of which resulted in higher ϕ than all other CMAES_P treatments when SVD or LSQR were also used. This was most likely because covariance matrix adaption strategy was developed using the principal of maximum entropy, and this principal is enacted by minimising the amount of prior information built into the distribution (Hansen and Ostermeier, 2001; Jaynes, 1982). Having *a priori* information implicit to the CMAES_P treatments 66-68 through the form of initial parameter estimates appears

to have biased the covariance matrix away from evolution towards the parameter vector providing the global minimum.

As we did not alter either the weighting associated with parameters in prior information equations, or the overall weighting associated with prior information articles, it is possible that different parameter factors or article weighting may have improved the accuracy and precision of the modelled data. Provision of different initial parameter estimates and weighting for both parameters and equations is an exercise that remains to be conducted for both CMAES_P and Tikhonov optimisation algorithms.

Certain control file settings of CMAES_P allowed reduction of ϕ to lower values than those obtained by using Tikhonov regularisation (cf. Tables 2 and 3). Convergence of the gradient-based methods with regularisation to ϕ values of between 26 and 27 indicates that these algorithms terminated on local minima, compared with some of the CMAES_P methods, which reduced ϕ further, as low as 15.9 (Table 7). Given that three CMAES_P treatments converged on the same ϕ value, (treatments 61, 65 and 69), it is possible that this value represents the global minimum of the solution. Our results also suggest that treatments with Tikhonov regularisation converged on local minima related to the N_2O emissions measurements (Fig. 2), since this variable primarily contributed to ϕ in regularisation treatments, in contrast to treatments with CMAES_P that were more effective at lowering the contribution to ϕ caused by N_2O emissions.

Optimisation settings resulting in low model accuracy and precision

Without Tikhonov regularisation, optimisation performance by the default algorithm was improved via reduced number of model calls when composite parameter sensitivity reuse was implemented (see Table 5). However, employing SENREUSE in combination with regularisation (Table 6) generally degraded with quality of the data fit (cf. treatments 43 to 44, 39 to 40 or 45 to 46). Reusing composite parameter sensitivity from one iteration to the next clearly saved CPU time in re-evaluating sensitivity from one run to the next, however this reduction in computational time came at the expense of early convergence and higher terminal ϕ values, indicating the PEST setting SENREUSE should be avoided.

There were several other treatments that had little effect on the value of ϕ or r but required significantly longer to converge (Table 5). Implementing PEST's method to fit derivatives using five points (ALWAYS 5, treatment 15) caused significantly greater run time as opposed to the default forward differencing/three-point derivative computation. Increasing the tolerances on parameter convergence (i.e. increasing the number of iterations over which parameters do not change and the relative change from iteration to iteration; treatments 22 and 23) also resulted much longer run times but no improvement in model to measurement mismatch.

Model features enabling optimisation by PEST and applicability of our results to other studies

The use of PEST as an optimisation tool has several advantages. PEST (1) facilitates simultaneous parameter optimisation across multiple objective functions, (2) requires very little programming knowledge (control, instruction and template files are arranged in a straightforward and intuitive layout) and (3) contains several free utility programs within the suite of PEST tools that help create, guide and error-check files prior to parameterisation (e.g. utility programs for adding Tikhonov regularisation, replacing parameter values between one optimisation and the next, checking the

value of the Jacobian matrix from iteration to iteration, etc.). PEST also includes linear and nonlinear methodologies for quantification of predictive uncertainty (the software can be run in 'predictive analysis' or 'Pareto' modes, both of which facilitate pre- and post-calibration uncertainty analyses), although these features were not applied here in order to keep the study size manageable.

As mentioned above, some of the model features allowing optimisation by PEST include the availability of model file(s) with all of the parameters used in the model simulation (in this case these were APSIM .sim files, generated with the ApsimToSim executable), as well as corresponding model outputs (in this case APSIM .out files). The format(s) of model outputs must also be sufficiently consistent from simulation to simulation such that their locations can be uniquely specified in the PEST instruction files (.ins). Even so, PEST has considerable flexibility in its ability to recognise a specific model output (e.g. a model output measured on a specific date) and thus the change in modelled values for each parameter upgrade iteration. PEST also requires a "model command line" wherein a command line interpreter code (or equivalent) can be used to specify the location and run the model (e.g. for Treatment 5 we used "C:\Program Files (x86)\Apsim78-r3867\Model\ApsimModel.exe" PEST_treatment_5.sim > null", where the line in quotation marks provides the location of the ApsimModel executable, "PEST_treatment_5.sim" is the name of the APSIM .sim file containing PEST demarked parameters, and "> null" prevents run time information from APSIM showing on the command line). Thus, availability of a compiled version of the model allowing execution from the command line is a key feature facilitating optimisation by PEST.

We postulate that the results found in this study would also be applicable and relatively consistent for other models. Although models vary widely in function, intent and complexity, PEST was specifically designed to be model independent and thus applicable across a range of models. In general terms, use of Tikhonov regularisation with prior equations provided better fits than the standard GML algorithm, as did inclusion of SVD or LSQR (though regularisation treatments with LSQR were no necessarily faster than those with SVD). Combining SVD or LSQR with Tikhonov regularisation, log transforming parameters prior to inversion and setting the Marquardt lambda modification factor (RLAMFAC) to -3 further improved the quality of the optimisation obtained. Even less error variance resulted from using CMAES_P, especially with larger population sizes, albeit CMAES_P required much longer run time than the GML algorithm with Tikhonov regularisation. We cannot disqualify the assertion that any optimisation algorithm applied here resulted in overfitting, though regularisation *per se* is designed to guard against overfitting by allowing heterogeneity to emerge where its existence is supported by field data, and penalising regions of the solution space where model heterogeneity cannot be supported by the calibration data (Tonkin and Doherty, 2005). An opportunity for future studies would be to test our hypothesis that these results (i.e. order of PEST algorithms resulting in the best fit of simulations to measured data) is reasonably consistent across models by applying PEST to their own models under a diverse range of experimental conditions.

Limitations of this study

This study investigated how different PEST algorithms and control file settings influenced the value of ϕ, r and associated CPU time required for each optimisation treatment. For the default GML algorithm and that with Tikhonov regularisation, we fitted 117 measurements with 115 parameters. Various rules of thumb in statistics exist for the desirable ratio of measurements to parameters, but often a ratio of 10 is used (Harrell Jr. et al., 1984). This study was far from that. Further, some of model errors were not independent for different measurements, so the effective sample size would be less than 117. It also should be noted that this study was applied to data from a single location and treatment. In a general calibration-validation experiment that aimed to use the calibrated model

in a validation or to make predictions or scenario analysis such as genotype by environment by management analysis, the approach conducted here would not be appropriate because some of the treatments could potentially be overfitted. As well, conclusions regarding preferential treatments may be specific to this study and different conclusions as to the best calibration approach may apply to other situations. Nevertheless, the purpose of this study was not to use the calibrated model to make agronomic predictions. Rather, this study was designed to test the ability of PEST to optimise extreme ratios of number of parameters to field measurements. Indeed, the purpose of this work was to determine which PEST algorithms resulted in the best fit to the 13 datasets and thus provide guidance for future studies using PEST on the pros and cons of each of the main approaches, as well as the motivation for using each approach. For example, if CPU time is not an issue due to parallel computing through clusters, CMAES_P with increased population size and log transformation of variables prior to inversion would be recommended. In the case that optimisation runs are constrained by CPU time, Tikhonov regularisation with SVD and LSQR and log transformation is suggested.

For future studies that aim to use PEST for parameterisation of their model followed by model evaluation and/or application, we advise that model practitioners select a small number of the optimisation treatments shown here (e.g., four or five treatments with the lowest ϕ), calibrate their model on multiple datasets and experimental treatments (preferably with more than one location if the study is on an agronomic experiment similar to the present study), validation of their model (spatially and temporally if possible) and restriction on the number of parameters optimised such that the ratio of data points to optimised parameters is 10 or greater. Consideration of which parameters are optimised also needs attention. Here we identified parameters for optimisation through manual sensitivity analysis, wherein the magnitude of change in model outputs resulting from changes in single parameters was recorded. Future studies could automate this step by using some of the tools provided with PEST or related programs, either the PEST utility program SENSAN (which allows users to conduct local sensitivity analyses), or the global sensitivity analyser (contained in the PEST compatible program 'PEST++'; Welter et al., 2015). Selection of sensitive model parameters influencing the model output variable of interest is an important first step, particularly if the number of parameters to be calibrated is low, for selection of insensitive parameters (or sensitive model parameters with respect to another model output but not the output that is to be fitted) will likely limit the quality of model fit obtained and thus predictive skill of simulations conducted using the calibrated model. In this study, optimised parameters varied in function; some having impact on simulated crop biomass, others on soil water content, others on nitrous oxide emissions, etc. For crop models like APSIM, it is unlikely that one model parameter will have significant influence on all model outputs, meaning that more parameters will need to be chosen for optimisation if the number of variables fitted is greater.

Conclusions

This study demonstrated several approaches for automated parameterisation of the complex deterministic model, APSIM. Through use of the model-independent Parameter ESTimation software (PEST), we examined multi-objective parameterisation of APSIM using data from a maize cropping experiment that contained several datasets. Tikhonov regularisation generally improved the performance of the default Gauss-Levenberg-Marquardt algorithm, particularly when the scaling factor used to determine parameter increments in successive optimisation iterations (RLAMFAC) was set to a value of -3 (enabling PEST to scale λ during each iteration of the inversion process so that λ can achieve a value of 1.0 with three adjustments), and either LSQR decomposition or SVD was used. Nonetheless, employing the default GML algorithm with LSQR or SVD also resulted in high quality

calibration and in significantly less computational time than other optimisation algorithms examined here. Employing CMAES_P with log transformed parameters and increased population size resulted in very low ϕ values but required significantly longer to converge. We propose that auto-parameterisation could be used as a protocol in future model inter-comparison exercises, since it would (1) foster removal of some of the subjectivity in simulation results associated with anthropogenic parameterisation after predefining lower and upper parameter bounds, (2) allow standardisation of parameter convergence criteria within given optimisation runs, (3) accelerate and systematise the inverse modelling process (Necpálová et al., 2015), and (4) highlight important effects of structural differences between models.

Acknowledgements

Author contributions: MTH conceived and conducted the study, PPR and LZ supplied agronomic data; all authors contributed to writing the paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Akponikpè, P.B.I., Gérard, B., Michels, K., Bielders, C., 2010. Use of the APSIM model in long term simulation to support decision making regarding nitrogen management for pearl millet in the Sahel. European Journal of Agronomy 32(2) 144-154.

Alluvione, F., Bertora, C., Zavattaro, L., Grignani, C., 2010. Nitrous Oxide and Carbon Dioxide Emissions Following Green Manure and Compost Fertilization in Corn All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher. Soil Science Society of America Journal 74 384-395.

Alluvione, F., Fiorentino, N., Bertora, C., Zavattaro, L., Fagnano, M., Chiarandà, F.Q., Grignani, C., 2013. Short-term crop and soil response to C-friendly strategies in two contrasting environments. European Journal of Agronomy 45 114-123.

Alluvione, F., Halvorson, A.D., Del Grosso, S.J., 2009. Nitrogen, tillage, and crop rotation effects on carbon dioxide and methane fluxes from irrigated cropping systems. J Environ Qual 38(5) 2023-2033.

Archontoulis, S.V., Miguez, F.E., Moore, K.J., 2014. A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: Application to soybean. Environmental Modelling & Software 62 465-477.

Aster, R.C., Borchers, B., Thurber, C.H., 2013. Chapter Three - Rank Deficiency and Ill-Conditioning, Parameter Estimation and Inverse Problems (Second Edition). Academic Press: Boston, pp. 55-91.

Beck, M.B., Halfon, E., 1991. Uncertainty, identifiability and the propagation of prediction errors: A case study of lake ontario. Journal of Forecasting 10(1‐2) 135-161.

Buis, S., Wallach, D., Guillaume, S., Varella, H., Lecharpentier, P., Launay, M., Guérif, M., Bergez, J.-E., Justes, E., 2011. The STICS crop model and associated software for analysis, parameterization, and evaluation, In: (ed.), L.R.A.a.L.M. (Ed.), Methods of introducing system models into agricultural research. Advances in Agricultural Systems Modeling Series 2.: Avignon, France.

C4SF, 2017. C4 Site Factory. PEST model independent parameter estimation and uncertainty analysis. Regularization. Available http://www.pesthomepage.org/PEST_FAQ.php [Accessed 18 December 2018].

Carberry, P.S., Ramganathan, R., Reddy, L.J., Chauhan, Y.S., Robertson, M.J., 2001. Predicting growth and development of pigeonpea: flowering response to photoperiod. Field Crops Res. 43 1501-1513.

Chen, C., Berger, J., Fletcher, A., Lawes, R., Robertson, M., 2016. Genotype×environment interactions for phenological adaptation in narrow-leafed lupin: A simulation study with a parameter optimized model. Field Crops Research 197 28-38.

Chisanga, C.B., Phiri, E., Shepande, C., Sichingabula, H., 2015. Evaluating CERES-Maize model using planting dates and nitrogen fertilizer in Zambia. J. Agric. Sci. 7(3) 79-97.

Doherty, J., 2016a. PEST: model-independent parameter estimation user manual part I: PEST, SENSAN and global optimisers. 6th Edn. Watermark Numerical Computing.

Doherty, J., 2016b. PEST: model-independent parameter estimation user manual part II: PEST utility support software. 6th Edn. Watermark Numerical Computing.

Doherty, J., Hunt, R.J., 2009. Two statistics for evaluating parameter identifiability and error reduction. Journal of Hydrology 366(1) 119-127.

Doherty, J.E., Hunt, R.J., 2010. Approaches to highly parameterized inversion: a guide to using PEST for groundwater-model calibration: U.S. Geological Survey Scientific Investigations Report 2010–5169, 59 p.

Droogers, P., Immerzeel, W.W., Lorite, I.J., 2010. Estimating actual irrigation application by remotely sensed evapotranspiration observations. Agricultural Water Management 97(9) 1351-1359.

Ehrhardt, F., Soussana, J.F., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., Sándor, R., Smith, P., Snow, V., Migliorati, M.d.A., Basso, B., Bhatia, A., Brilli, L., Doltra, J., Dorich, C.D., Doro, L., Fitton, N., Giacomini, S.J., Grant, B., Harrison, M.T., Jones, S.K., Kirschbaum, M.U.F., Klumpp, K., Laville, P., Léonard, J., Liebig, M., Lieffering, M., Martin, R., Massad, R.S., Meier, E., Merbold, L., Moore, A.D., Myrgiotis, V., Newton, P., Pattey, E., Rolinski, S., Sharp, J., Smith, W.N., Wu, L., Zhang, Q., 2018. Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N2O emissions. Global Change Biology 24(2) e603-e616.

Farré, I., Robertson, M.J., Asseng, S., French, R.J., Dracup, M., 2004. Simulating lupin development, growth, and yield in a Mediterranean environment. Australian Journal of Agricultural Research 55(8) 863-877.

Field, J.L., Marx, E., Easter, M., Adler, P.R., Paustian, K., 2016. Ecosystem model parameterization and adaptation for sustainable cellulosic biofuel landscape design. GCB Bioenergy 8(6) 1106-1123.

Finsterle, S., Zhang, Y., 2011. Solving iTOUGH2 simulation and optimization problems using the PEST protocol. Environmental Modelling & Software 26(7) 959-968.

Gaucherel, C., Campillo, F., Misson, L., Guiot, J., Boreux, J.J., 2008. Parameterization of a process-based tree-growth model: Comparison of optimization, MCMC and Particle Filtering algorithms. Environmental Modelling & Software 23(10) 1280-1288.

Grignani, C., Alluvione, F., Bertora, C., Zavattaro, L., Fagnano, M., Fiorentino, N., Chiarandà, F.Q., Amato, M., Lupo, F., Bochicchio, R., 2012. Field Plots and Crop Yields Under Innovative Methods of Carbon Sequestration in Soil, In: Piccolo, A. (Ed.), Carbon Sequestration in Agricultural Soils: A Multidisciplinary Approach to Innovative Methods. Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 39-60.

Hansen, N., 2016. The CMA Evolution Strategy: A Tutorial. ArXiv e-prints, arXiv:1604.00772. pp.1-39. 2005. <hal-01297037>.

Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. .Evolutionary Computation 9(2) 159–195.

Harrell Jr., F.E., Lee, K.L., Califf, R.M., Pryor, D.B., Rosati, R.A., 1984. Regression modelling strategies for improved prognostic prediction. Statistics in Medicine 3(2) 143-152.

Harrison, M.T., Evans, J.R., Moore, A.D., 2012. Using a mathematical framework to examine physiological changes in winter wheat after livestock grazing: 1. Model derivation and coefficient calibration. Field Crops Research 136(0) 116-126.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62 327-350.

Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom, G., White, J.W., Thorburn, P., 2015. Agricultural production systems modelling and software: Current status and future prospects. Environmental Modelling & Software 72 276-286.

Hunt, L.A., Pararajasingham, S., Jones, J.W., Hoogenboom, G., Imamura, D.T., Ogoshi, R.M., 1993. GENCALC: Software to facilitate the use of crop models for analyzing field experiments. Agronomy Journal 85(5) 1090-1094.

James, T.M., John, D., 2005. A hybrid regularized inversion methodology for highly parameterized environmental models. Water Resources Research 41(10).

Jaynes, E.T., 1982. On the rationale of maximum-entropy methods. Proceedings of the IEEE 70(9) 939-952.

Jones, J.W., He, J., Boote, K.J., Wilkens, P., Porter, C.H., Hu, Z., 2011. Estimating DSSAT cropping system cultivar-specific parameters using Bayesian techniques. Methods of Introducing System Models into Agricultural Research. Advances in Agricultural Systems Modeling 2 365-394.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. European Journal of Agronomy 18(3) 235-265.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18(3-4) 267-288.

Lacroix, M.P., Martz, L.W., Kite, G.W., Garbrecht, J., 2002. Using digital terrain analysis modeling techniques for the parameterization of a hydrologic model. Environmental Modelling & Software 17(2) 125-134.

Lampe, M., Willenbockel, D., Ahammad, H., Blanc, E., Cai, Y., Calvin, K., Fujimori, S., Hasegawa, T., Havlik, P., Heyhoe, E., Kyle, P., Lotze‐Campen, H., d'Croz, D.M., Nelson, G.C., Sands, R.D., Schmitz, C., Tabeau, A., Valin, H., Mensbrugghe, D., Meijl, H., 2014. Why do global long‐term scenarios for agriculture differ? An overview of the AgMIP Global Economic Model Intercomparison. Agricultural Economics 45(1) 3-20.

Lanczos, C., 1961. Linear Differential Operators, Chapter 3. Van Norstrand, London.

Lin, Y., O'Malley, D., Vesselinov, V.V., 2016. A computationally efficient parallel Levenberg-Marquardt algorithm for highly parameterized inverse model analyses. Water Resources Research 52(9) 6948-6977.

Lin, Z., 2005. Theory for PEST users. Department of Crop and Soil Sciences University of Georgia, Athens, GA 30602. Available: https://www.researchgate.net/publication/229004529_Theory_for_PEST_Users [Accessed 19 December 2018].

Marin, F.R., Jones, J.W., Royce, F., Suguitani, C., Donzeli, J.L., Filho, W.J.P., Nassif, D.S.P., 2011. Parameterization and Evaluation of Predictions of DSSAT/CANEGRO for Brazilian Sugarcane. Agronomy Journal 103(2) 304-315.

Marquardt, D.W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics 11(2) 431-441.

Moore, C., Doherty, J., 2006. The cost of uniqueness in groundwater model calibration. Advances in Water Resources 29(4) 605-623.

NAG, 1983. Numerical Algorithms Group. The NAG Fortran PC50 Library Handbook - release 1.

Necpálová, M., Anex, R.P., Fienen, M.N., Del Grosso, S.J., Castellano, M.J., Sawyer, J.E., Iqbal, J., Pantoja, J.L., Barker, D.W., 2015. Understanding the DayCent model: Calibration, sensitivity, and identifiability through inverse modeling. Environmental Modelling & Software 66 110-130.

Paige, C.C., Saunders, M.A., 1982. LSQR - An algorithm for spare linear equations and sparse least squares. Acm Transactions on Mathematical Software 8(1) 43-71.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proceedings of the National Academy of Sciences 111(9) 3268-3273.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agricultural and Forest Meteorology 170 166-182.

Rouchier, S., Woloszyn, M., Kedowide, Y., Bejat, T., 2015. Identification of the hygrothermal properties of a building envelope material by the covariance matrix adaptation evolution strategy. Journal of Building Performance Simulation.

Samanta, S., Mackay, D.S., 2003. Flexible automated parameterization of hydrologic models using fuzzy logic. Water Resources Research 39(1) SWC 1-1-SWC 1-13.

Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go? European Journal of Agronomy 94 25-35.

Sequeira, R.A., Olson, R.L., Willers, J.L., McKinion, J.M., 1994. Automating the parameterization of mathematical models using genetic algorithms. Computers and Electronics in Agriculture 11(2) 265-290.

Sexton, J., Everingham, Y., Inman-Bamber, G., 2016. A theoretical and real world evaluation of two Bayesian techniques for the calibration of variety parameters in a sugarcane crop model. Environmental Modelling & Software 83 126-142.

Tonkin, M.J., Doherty, J., 2005. A hybrid regularized inversion methodology for highly parameterized environmental models. Water Resources Research 41(10).

Turpin, J.E., Robertson, M.J., Haire, C., Bellotti, W.D., Moore, A.D., Rose, I., 2003. Simulating fababean development, growth, and yield in Australia. Australian Journal of Agricultural Research 54(1) 39-52.

Venterea, R.T., Baker, J.M., 2008. Effects of Soil Physical Nonuniformity on Chamber-Based Gas Flux Estimates All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher. Soil Science Society of America Journal 72(5) 1410-1417.

Wallach, D., Keussayan, N., Brun, F., Lacroix, B., Bergez, J.-E., 2012. Assessing the uncertainty when using a model to compare irrigation strategies. Agronomy Journal 104(5) 1274-1283.

Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2014. Chapter 6 - Parameter Estimation with Classical Methods (Model Calibration), In: Wallach, D., Makowski, D., Jones, J.W., Brun, F. (Eds.), Working with Dynamic Crop Models (Second Edition). Academic Press: San Diego, pp. 205-276.

Welter, D.E., White, J.T., Hunt, R.J., Doherty, J.E., 2015. Approaches in highly parameterized inversion—PEST++ Version 3, a Parameter ESTimation and uncertainty analysis software suite optimized for large environmental models, Techniques and Methods: Reston, VA, p. 64.